

The evolution of altruistic defense traits in structured populations

Dissertation der Fakultät für Biologie
der Ludwig-Maximilians-Universität München
zur Erlangung des Doktorgrades der Naturwissenschaften

vorgelegt von
Felix Jordan
aus Wiesbaden

München, den 29. September 2015

Erstgutachter: Prof. Dr. Dirk Metzler
Zweitgutachter: Prof. Dr. Niels Dingemanse

Tag der Abgabe: 29.09.2015
Tag der mündlichen Prüfung: 26.01.2016

Erklärung

Diese Dissertation wurde im Sinne von §12 der Promotionsordnung von Prof. Dr. Dirk Metzler betreut. Ich erkläre hiermit, dass die Dissertation nicht einer anderen Prüfungskommission vorgelegt worden ist und dass ich mich nicht anderweitig einer Doktorprüfung unterzogen habe.

Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist.

Im Abschnitt Contribution ist mein Anteil an den einzelnen Projekten dargestellt.

München, den 29. September 2015

Felix Jordan

Contents

Summary	6
Zusammenfassung	8
Contribution	10
1 Introduction	11
1.1 Altruism	11
1.1.1 The puzzle of altruism	11
1.1.2 Altruism in nature	12
1.1.3 Explanations of the evolution of altruism	12
1.2 Altruistic defense	20
1.2.1 Parasitism	20
1.2.2 Altruistic defense	21
1.2.3 The slavemaking ant <i>Protomognathus americanus</i> and its hosts . .	22
1.2.4 Slave rebellion	24
1.3 Goals and outline	25
2 Models and methods	27
2.1 Mathematical analysis and asymptotics	27
2.1.1 Individual-based model	27
2.1.2 Diffusion approximation	29
2.1.3 Diffusion model	29
2.1.4 Convergence of the relative frequency of altruists	30
2.1.5 McKean-Vlasov limit	31
2.1.6 Long-term behavior of the average altruist frequency	32
2.1.7 Invasion of an altruistic defense allele	32
2.2 Simulations on the evolution of altruistic defense	33

2.2.1	Small populations	33
2.2.2	Large populations	38
2.3	Evolution of slave rebellion in ants	41
3	Results	57
3.1	Mathematical analysis and asymptotics	57
3.2	Simulations on the evolution of altruistic defense	62
3.2.1	Small populations	62
3.2.2	Large populations	65
3.3	Evolution of slave rebellion in ants	72
4	Proofs	80
4.1	Mathematical analysis and asymptotics	80
4.1.1	Notation	80
4.1.2	Diffusion approximation	80
4.1.3	Diffusion model	85
4.1.4	Convergence of the relative frequency of altruists	86
4.1.5	McKean-Vlasov limit	129
4.1.6	Long-term behavior of the average altruist frequency	132
4.1.7	Invasion of an altruistic defense allele	137
5	Discussion	141
	Bibliography	161
	Acknowledgments	175

Summary

An altruistic trait reduces the reproductive success of its bearer, while at the same time increasing the reproductive success of other individuals in the population. If this trait is aimed at defending against a parasite or predator, it is called an altruistic defense trait. The act of defending entails a cost to the individual and helps the population by reducing the pressure from parasite or predator. Determining whether such a costly defense mechanism can evolve is not trivial, as the benefit is not provided immediately, but is mediated by the parasite or predator population.

In this dissertation, a stochastic model is introduced to study the evolution of an altruistic defense trait. The host and parasite (or, equivalently, prey and predator) populations are structured in demes, that are connected by migration. There are two types of host (prey) individuals—one has a reproductive disadvantage compared to the other but locally reduces the parasite (predator) growth rate. The interactions are described by a birth-death process, for which the diffusion approximation is computed. It takes the form of stochastic Lotka-Volterra equations. For these, the relative frequency of altruists is shown to converge to spatially structured Wright-Fisher diffusions in the limit of infinitely large population sizes. Then, the many-demes limit of these diffusions is computed and a condition for extinction or fixation of altruists, depending on the ecological model parameters, is derived. Furthermore, the invasion process is considered. In a non-altruistic population, altruists arise in small positive frequency in a single deme and a condition for survival is obtained.

The robustness of the theoretical prediction regarding the success of an altruistic defense allele is tested with computer simulations using two different models. First, the individual-based model that is the starting point of the theoretical analysis is simulated. Here, finite population sizes are assumed. Individuals produce offspring, die, or migrate at specified rates. The other simulation model approximates the limit process describing the relative frequency of altruists, as obtained in the theoretical analysis. For this, infinite population sizes and a complete separation of ecological and evolutionary time scales are assumed. Thus, with these two simulation models, the effect of assuming an infinite population size compared to a finite one can be investigated together with the consequences of a separation

of time scales. Furthermore, some extensions of the simplified model are investigated with the simulation approaches, i.e., assuming random fluctuations of the competition strengths within species as well as varying the migration schemes.

Finally, the concrete case of slave rebellion in ants is investigated. A detailed model of interactions between slavemaking ants and their hosts on a patch of forest is defined and studied with extensive computer simulations. First, field data from New York State and West Virginia is used in an Approximate Bayesian Computation approach to find plausible parameter settings. With these, several simulation series are performed to understand the mechanisms that are relevant for the evolution of a hypothetical allele for slave rebellion. A metapopulation of several connected patches is simulated to assess whether a benefit on a larger spatial scale may influence the evolution of slave rebellion.

Zusammenfassung

Ein altruistisches Merkmal reduziert den Fortpflanzungserfolg seines Trägers, während es gleichzeitig den Fortpflanzungserfolg anderer Individuen in der Population steigert. Wenn dieses Merkmal der Verteidigung gegen einen Parasiten oder Räuber dient, spricht man von altruistischer Verteidigung. Der Akt der Verteidigung geht zu Kosten des Individuums und hilft der Population, indem es den Druck durch Parasiten oder Räuber senkt. Zu bestimmen, ob solch ein kostenbehafteter Verteidigungsmechanismus evolvieren kann, ist nicht trivial, da der Vorteil nicht direkt zum Tragen kommt, sondern mittelbar durch den Parasiten bzw. Räuber wirkt.

In dieser Dissertation wird ein stochastisches Modell eingeführt, um die Evolution eines altruistischen Verteidigungsmerkmals zu studieren. Die Wirte und Parasiten (bzw. Beute und Räuber) sind in Subpopulationen strukturiert, die durch Migration verbunden sind. Es gibt zwei Typen von Wirts- (Beute-) Individuen. Ein Typ hat, verglichen zum anderen, einen Nachteil in der Fortpflanzung, aber reduziert lokal die Wachstumsrate der Parasiten (Räuber). Die Interaktionen sind durch einen Geburts-Todes-Prozess beschrieben, für welchen die Diffusionsapproximation berechnet wird. Sie hat die Form stochastischer Lotka-Volterra-Gleichungen. Für diese wird gezeigt, dass die relative Häufigkeit der Altruisten im Grenzwert von unendlichen Populationsgrößen gegen räumlich strukturierte Wright-Fisher Diffusionen konvergiert. Dann wird der Grenzprozess für unendlich viele Subpopulationen bestimmt und eine Bedingung für Aussterben oder Fixierung der Altruisten, in Abhängigkeit der ökologischen Modellparameter, wird hergeleitet. Außerdem wird der Invasionsprozess betrachtet. In einer nicht-altruistischen Population treten Altruisten in kleiner positiver Frequenz in einer einzigen Subpopulation auf, und eine Bedingung für ihr Überleben wird ermittelt.

Die Robustheit der theoretischen Resultate im Hinblick auf den Erfolg eines Allels für altruistische Verteidigung wird mit Computersimulationen anhand zweier unterschiedlicher Modelle getestet. Zunächst wird das individuenbasierte Modell simuliert, das den Ausgangspunkt der theoretischen Analyse darstellt. Hier werden endliche Populationsgrößen angenommen. Individuen produzieren Nachwuchs, sterben oder migrieren mit festgelegten Ra-

ten. Das andere Simulationsmodell approximiert den in der theoretischen Analyse ermittelten Grenzprozess, der die relative Häufigkeit der Altruisten beschreibt. Hier werden unendliche Populationsgrößen und eine komplette Trennung der ökologischen und evolutionären Zeitskalen angenommen. Daher kann mit diesen beiden Simulationsmodellen der Effekt der Annahme unendlich großer Populationen im Vergleich zu endlich großen zusammen mit den Folgen einer Trennung der Zeitskalen untersucht werden. Außerdem werden einige Erweiterungen des vereinfachten Modells mit Hilfe der Simulationsansätze untersucht, indem zufällige Fluktuationen des Konkurrenzdrucks innerhalb der Spezies angenommen und die Migrationsmuster variiert werden.

Schließlich wird der konkrete Fall von Sklavenrebellion bei Ameisen betrachtet. Ein detailliertes Modell für Interaktionen von Sklavenhalterameisen und ihren Wirten in einem Waldgebiet wird definiert und mit umfassenden Computer-Simulationen studiert. Zunächst werden Feld-Daten aus den amerikanischen Bundesstaaten New York und West Virginia genutzt, um plausible Parameter-Einstellungen zu bestimmen. Anhand dieser werden mehrere Simulationsserien durchgeführt, um Mechanismen und Parameter zu finden, die relevant sind für die Evolution eines hypothetischen Allels für Sklavenrebellion. Eine Metapopulation mehrerer verbundener Gebiete wird simuliert, um zu untersuchen, ob ein Vorteil auf einer größeren räumlichen Skala die Evolution der Sklavenrebellion beeinflussen kann.

Contribution

The initial idea for the mathematical analysis came from Martin Hutzenthaler and Dirk Metzler. I completed all of the proofs with the help of Martin Hutzenthaler and I wrote the manuscript (Hutzenthaler et al., 2015), except for the introduction, which was mainly written by Martin Hutzenthaler, with help by Dirk Metzler and me.

The simulation studies of the second part were outlined by Dirk Metzler and Martin Hutzenthaler. I wrote the simulation programs, performed the simulation studies, did the data analyses, and interpreted the results.

For the simulations of ant interactions, I added several features to the simulation program. These features include the simulation of genetic information of host ants, storage of genetic data for slaves in slavemaker nests, the possibility of polygynous host nests, and a metapopulation structure. The simulation studies were designed by Dirk Metzler and performed by me. Furthermore, I wrote the scripts to compute all summary statistics and relatedness values, performed the ABC analysis, and pre-processed the output of the simulations. The data was mainly analyzed by Dirk Metzler with my help. Dirk Metzler also wrote most of the manuscript (Metzler et al., 2015). I wrote the description of the simulation program.

Chapter 1

Introduction

1.1 Altruism

1.1.1 The puzzle of altruism

In his famous book *The origin of species*, Darwin (1859) introduced the theory of evolution via natural selection. Although being generally accepted, at first it remained a challenge to explain why there exist self-sacrificial behaviors in nature. Despite Darwin (1859, Chapter 8) noting that “selection may be applied to the family, as well as to the individual”, the possibility of an individual disadvantage being overcome by an advantage to the group or even the species was generally rejected (see Fisher, 1958 and Wright, 1948, as well as Hamilton, 1963 for a brief overview). When viewing the individual as the agent that natural selection acts upon, it seems impossible that any character could evolve that harms the reproductive success of its bearer. If such a character at the same time increases the reproductive success of other individuals, then that character is called altruistic. Sometimes this is also referred to as strong altruism. On the other hand, weak altruism describes an action that entails a relative decrease in the focal individual’s reproductive success compared to its interaction partners, although in absolute terms it might be beneficial for the actor (Wilson, 1979).

At first sight, it seems like non-altruistic individuals—often referred to as free-riders or cheaters—would have a selective advantage compared to altruists. These cheaters can exploit the helping of altruists without contributing any help themselves, and thereby not suffering any costs in terms of reproductive success. Hence, a cheater should on average have more descendants than an altruist and thus, altruists should over time be driven to extinction. That reasoning would also suggest that any altruistic trait should be kept from reaching high frequencies in any given population. Consequently, it should be highly improbable to find any form of cooperation in nature.

1.1.2 Altruism in nature

Despite the apparent disadvantage of altruistic traits, there are numerous variations of altruism occurring in nature. A famous example of altruistic behavior observed in nature is food sharing in vampire bats (Wilkinson, 1984, 1990). After two consecutive nights without a meal from blood of their prey, vampire bats die. Individuals that did not obtain food during one night are fed by others in their group and thus saved from dying of food deprivation. This is of course costly to the helper, who sacrifices a share of food and thereby risks earlier starvation. Another case of altruism is the production of costly public goods in bacteria (Diggle et al., 2007). Bacterial cells use diffusible signal molecules as a means to communicate with each other. This process—known as quorum sensing (Williams et al., 2007)—is open to exploitation. Members of the group can refuse to produce the signals themselves to save the associated production cost but instead completely rely on others to invest in the necessary communication. A different kind of altruistic behavior is shown by coatimundis living in small groups. They groom each other, which can provide the benefit of a reduction in the number of ectoparasites (Wilkinson, 1988). This social grooming infers the cost of time investment, which reduces time spent foraging and, probably more importantly, it reduces vigilance times, thereby increasing the risk of an attack (Wilkinson, 1988). An extreme example of altruism is that of sterility among social insects. In the phenomenon known as “eusociality”, certain individuals completely forgo reproduction to instead spend their time helping others and raising their offspring (e.g., Hölldobler and Wilson, 2009). Consequently, these individuals will not have any offspring to carry their genes to the next generation. This baffling case has already been discussed by Darwin (1859) and is still debated in recent literature (Nowak et al., 2010; Abbot et al., 2011)

1.1.3 Explanations of the evolution of altruism

Resolutions to the puzzling existence of altruism and explanations on how it may evolve are provided by the theories of inclusive fitness and group selection and by game theory. An overview of the different mechanisms that can explain the evolution of cooperative behavior is given by West et al. (2007a). One main idea is to incorporate spatial structure into the analysis, which may allow altruists to co-occur and help each other more than non-altruists.

1.1.3.1 Kin selection

The theory of inclusive fitness, which is also referred to as kin selection, is introduced by Hamilton (1963, 1964a,b). The idea is that the spread of a gene is not determined by whether

it is beneficial to its bearer but rather to the gene itself. Instead of looking at the direct fitness effect of the behavior on the actor, the evolutionary success of a gene can be determined by investigating the inclusive fitness, which also takes into account the effect that the behavior has on other individuals. By promoting the reproductive success of other individuals, a gene that is detrimental to its bearer may spread if the relatedness to the beneficiaries as well as the amount of benefit is large enough. This result is known as Hamilton's rule (Hamilton, 1963, 1964a,b), which, more precisely, states that an altruistic trait will spread if

$$r \cdot b > c, \quad (1.1)$$

where r is the relatedness of the actor to beneficiaries, b is the increase in reproductive success of the beneficiaries, and c is the loss in reproductive success of the actor. The relatedness r is strictly defined as a regression coefficient (Gardner et al., 2011, Box 5), but can be interpreted as a measurement of sharing a focal gene compared to the population average gene frequency (Bourke, 2011, Box 1). A high relatedness thus means a high chance of sharing a certain gene, which does not necessarily have to be caused by kinship. When determining b and c , it is important to only account for changes in reproductive success caused by the behavior in question. Actually, equation (1.1) is the abbreviated version of Hamilton's formula. The extended form says that an action is selectively favored if

$$\sum_{i \in \mathcal{I}} r_i \cdot b_i > 0, \quad (1.2)$$

where \mathcal{I} denotes the set of all individuals in the population, b_i denotes the effect in terms of reproductive success that the action of the focal individual has on any individual $i \in \mathcal{I}$ (including itself), and r_i denotes their relatedness (e.g., Grafen, 2007; Lion et al., 2011). Note that the effect in terms of reproductive success of performing the altruistic action is negative for the focal individual itself. Thereby, calling the focal individual i_0 , the term $-b_{i_0}$ is commonly referred to as "cost" c , with $c > 0$. Furthermore, since the relatedness of the focal individual to itself equals one, equation (1.2) can be rewritten as

$$\sum_{i \in \mathcal{I} \setminus i_0} r_i \cdot b_i > c, \quad (1.3)$$

to resemble equation (1.1). An alternative interpretation of Hamilton's rule is the "neighbor-modulated fitness", where fitness effects of all other individuals in the population on the reproductive success of the focal individual are captured in the b_i terms. This is equivalent to the inclusive fitness view described above (e.g., Gardner, 2008, Box 2). Hamilton's rule

not only predicts under which conditions altruism evolves, but it similarly also predicts if cooperation ($b > 0$, $c < 0$ in notation of equation (1.1)), selfishness ($b < 0$, $c < 0$), spite ($b < 0$, $c > 0$), or any other set-up of mixed positive and negative effects (using equation (1.2)) can evolve (Bourke, 2011).

1.1.3.2 Group selection

Another explanation of the evolution of altruism is provided by the theory of group selection, which captures the “process of genetic change which is caused by the differential extinction or proliferation of groups of organisms” (Wade, 1978). An overview of the development of the theory of group selection is given by Traulsen and Nowak (2006), who attribute the first mathematical model of group selection to Wright (1945). In group selection models, the relatedness between individuals is not calculated. Instead, it is assumed that the population is split into groups that are in direct competition with each other. The idea then is that groups with a large frequency of altruists have an advantage over groups with a low altruist frequency. A classical model of group selection is the haystack model proposed by Maynard Smith (1964), which describes a population of mice structured in separated haystacks. A single biallelic locus determines whether a mouse is timid or aggressive, with timidity being recessive. If a haystack contains both types, then the aggressive ones will drive out all timid mice from the haystack, such that each haystack consists of only timid or only aggressive mice when migration occurs. During migration, all haystacks are abandoned and there is random mating across the whole population. Then the haystacks are freshly populated, each by a single fertilized female. Timid groups are assumed to produce more migrants and a condition is derived under which this group-level advantage can override the individual disadvantage, such that timidity is selectively favored. The author himself deems this model too artificial to be further pursued (Maynard Smith, 1964). Nevertheless, it is an example of when group selection can predict the evolution of self-sterilizing behavior under restrictive conditions. A review of group selection models is given by Wade (1978). The models discussed in Wade (1978) make assumptions that disfavor the effect of group selection. They all show that only under very restrictive parameter settings, can a negative within-group selection be overcome by between-group selection. Since it is deemed unlikely that any natural population would adhere to such narrow parameter ranges, these studies conclude that group selection is not strong enough to be a driver of changes in gene frequencies on an evolutionary scale (see Wade, 1978, and the references therein). However, Wade (1976) experimentally shows that group selection can overcome individual selection for a population of *Tribolium castaneum*. Wade (1978) suggests that changing some of the assumptions that

are implicit in previous models might show an increased effect of group selection. One of the restricting assumptions is that of a “migrant pool”, that each deme contributes to in equal amount (Wade, 1978). From this migrant pool, colonists that form new subpopulations are drawn randomly. By introducing a “propagule pool” to which each deme sends groups of migrants (the propagules) and recolonizing empty sites with a randomly chosen propagule from that pool, Slatkin and Wade (1978) adapt classical models to allow for a larger effect of group selection on a quantitative trait. This is achieved by diminishing the stage of complete mixing among migrants, which is present in classical models using a migrant pool. Consequently, between-deme variance of the quantitative trait is increased and the effect of selection on the level of groups is stronger. The result of Slatkin and Wade (1978) rests on the assumption that the variance of the trait is the same in each deme. Although this is an idealization, they argue that it is reasonable if there is gene flow between demes, which reduces differences in variance. Furthermore, they refer to Lande (1976) who shows (for polygenic traits with linked loci) that mutation can enhance genetic variation. More generally, this result by Lande (1976) is important for all models of group selection, due to showing how variance between demes can be maintained (which is necessary for group selection to act), even if selective forces tend to reduce it. Thereby, group selection may be more relevant than suggested by several models, for which genetic drift and random sampling from the migrant pool are the only forces of maintaining genetic variation between demes (Wade, 1978).

Wade (1978) notes that the assumption of all demes sending off the same number of migrants restricts the possibility for group selection to act in numerous models. By allowing for empty sites within demes and determining the amount of migrants in dependence of the current size of a deme, Alizon and Taylor (2008) remove this assumption from their model and show that altruism can be favored. Furthermore, their approach reduces the impact of competition among altruists, which has been shown to negate cooperative benefits and prevent the evolution of altruism in other models (Wilson et al., 1992; Taylor, 1992).

The haystack model of Maynard Smith (1964) also makes the assumption of a migrant pool, from which individuals are randomly drawn to mate and start colonies at empty sites, but with the modification that a fixed proportion of migrating females mates with males from their own group. Thereby, global mixing in the migration stage is reduced and the strength of group selection is enhanced. Furthermore, Maynard Smith (1964) notes that times at which groups are reproductively isolated are necessary to ensure variation between groups. However, as Wade (1978) notes, this is merely a consequence of the extreme amount of individual selection assumed in the model, in which a single non-altruist will take over any group of altruists. Moreover, Wade (1978) states that although some reduction of population mixing is necessary for group selection to be important, between-deme variance can be maintained

even for moderate levels of migration, as shown by Wright (1931).

It is unclear how well the conditions in various models reflect the situation of natural populations (Maynard Smith, 1964). Yet, in a review of experimental studies of group selection, Goodnight and Stevens (1997) reckon that group selection may be more common than previously thought. Further models of group selection are developed by Van Valen (1971), Wilson (1975), and Traulsen and Nowak (2006). In the model by Traulsen and Nowak (2006), individuals in groups with many cooperators have a higher fitness. If the maximum group size is reached, a group can split in two, which will cause another group to be removed. Thereby, groups with many cooperators will split earlier than other groups, which can lead to a selective advantage of altruists under a simple condition of the model parameters. The theory of group selection can be seen as a special case of the more general idea of multilevel selection, which states that selection can act on different levels (Wade, 1982; Bijma et al., 2007; Gardner, 2015). These levels are not restricted to individuals and groups, but can also refer to families, species or genes. In particular, several levels may be involved simultaneously, as, e.g., in the scenario of interactions at the levels of individuals, families, and groups, investigated by Wade (1982). Therein, total gene frequency change of the allele in question is split into its components acting on the different levels. Selection on each level may work to increase or decrease the frequency of the allele, and therefore, Wade (1982) concludes, that all relevant levels have to be taken into consideration, when determining the success of the allele. Similarly, Bijma et al. (2007) show how to generalize the group selection approach. They partition the effect of selection into the different components, in the setting of individuals living in groups and groups forming metagroups. In particular, this approach of Bijma et al. (2007) can be applied to any number of levels of selection.

As pointed out by Gardner (2015), explaining the evolution of altruism by multilevel selection is actually an example of the more general Simpson's paradox, which is also known as the Yule-Simpson effect (Yule, 1903; Simpson, 1951; Blyth, 1972). This classical result from statistics describes how the aggregation of data from several groups can reverse the trend occurring within each group. In the case of altruism this would mean that within each group, there is a negative correlation between behaving altruistically and offspring numbers, while overall there can be a positive correlation. This is due to the fact that group sizes vary and larger (or faster-splitting) groups contribute more, when the population is viewed as a whole.

1.1.3.3 Connection of kin selection and group selection

A formal link between kin and group selection can be established using a very general concept of evolutionary change, the Price equation (Price, 1970; Marshall, 2011). This famous result partitions changes in gene frequencies across generations in changes due to selection and changes due to transmission (see also Gardner, 2008, 2015). Let w be the number of offspring of an individual in the population, let z be the individual's value of the character being investigated, such as behaving altruistically, denote their population averages by \bar{w} and \bar{z} , respectively, and define the change in character value between the individual and its offspring by Δz . Then the Price equation determines the change in population average of the character value, $\Delta \bar{z}$, to be

$$\Delta \bar{z} = \text{Cov} \left[\frac{w}{\bar{w}}, z \right] + \mathbb{E} \left[\frac{w}{\bar{w}} \Delta z \right]. \quad (1.4)$$

Hence, the causes for gene frequency changes can be separated. The first part denotes change due to selection and the second part denotes change due to transmission. The Price equation holds for any kind of ploidy, dominance, or epistasis, for sexual as well as asexual reproduction, and for random or deterministic mating (Price, 1970). Furthermore, applicability is not restricted to changes in gene frequencies, but also holds for any other type of entity that changes over time (Gardner, 2008). Variants of the Price equation for kin and group selection are derived by Queller (1992). Similarly, Marshall (2011) shows how Hamilton's rule and the approach of group selection theory can both be derived using the Price equation.

The relative importance of kin versus group selection is intensely debated in recent literature (Traulsen and Nowak, 2006; Lehmann et al., 2007; Fletcher and Zwick, 2007; Van Veelen, 2009; Traulsen, 2010; Nowak et al., 2010; Abbot et al., 2011; Marshall, 2011; Gardner et al., 2011). Van Veelen (2009) and Nowak et al. (2010) state that the inclusive fitness approach does not always work and can lead to wrong conclusions. Nowak et al. (2010) have provoked several responses supporting the theory of kin selection (Abbot et al., 2011; Boomsma et al., 2011; Strassmann et al., 2011; Ferriere and Michod, 2011; Herre and Wcislo, 2011) with a total of over 140 authors. Furthermore, Rousset and Lion (2011) point out that the main argument of Nowak et al. (2010) against inclusive fitness theory is based on a misinterpretation of the cost and benefit terms in Hamilton's rule, and conclude that it is invalid. Traulsen (2010) and Van Veelen et al. (2012, 2014) state that there are fundamental differences in the underlying theories of group and kin selection and that they cannot always be used interchangeably. On the other hand, Wade (1980) and Marshall (2011) derive the equivalence between the two approaches. Birch and Okasha (2015) clarify that the contradictions regarding the validity of kin selection theory stem from the fact that there are three

different formulations and interpretations of Hamilton's rule. First, there are approaches for which cost and benefit are defined as payoff parameters of the particular model (e.g., Nowak et al., 2010; Van Veelen et al., 2012). These definitions require some restrictions, such as an additive payoff structure, for Hamilton's rule to apply (Birch and Okasha, 2015). Second, in some approaches cost and benefit terms are partial regression coefficients, determining how the focal individual's genotype (or phenotype) is associated with its fitness and the fitness of its interaction partners (e.g., Queller, 1992; Gardner et al., 2011). This is the most general formulation and it is true for any model because it makes no assumptions about how the statistical relationships arise (Birch and Okasha, 2015). Third, other approaches use first-order approximations of the general formulation by replacing the regression coefficients with partial derivatives of a fitness function (e.g., Taylor and Frank, 1996; Lehmann and Rousset, 2010). These approximations are valid only if selection is weak and fitness effects are additive. Hence, the contradicting views on the validity and generality of Hamilton's rule stem from the use of different approaches (Birch and Okasha, 2015), meaning that Hamilton's rule always holds in its most general form, but that special care needs to be taken considering the assumptions, when using the other two approaches. The critiques against inclusive fitness theory are reviewed by Bourke (2011), who concludes that none of the advances is effective and, consequently, that inclusive fitness theory is deservedly the "leading theory for social evolution". West et al. (2007b) discuss some confusions in the usage of different terms related to altruism, kin and group selection, and state that these semantical issues unnecessarily hinder the progress of the field. Okasha (2010) criticizes the debate as a whole for demanding too many resources, because both approaches—kin and group selection—have provided valuable insights and merely offer different ways of understanding the evolution of social behavior. Similarly, acknowledging the value of both approaches, Lion et al. (2011) argue that most of the debate stems from implicit differences in the simplified models and that further discussion of these differences is of little help when studying evolutionary change in nature.

1.1.3.4 Game theory

Further important insights in understanding the evolution of altruistic behavior have been provided by evolutionary game theory (Axelrod and Hamilton, 1981; Nowak and May, 1992; Lieberman et al., 2005; Ohtsuki et al., 2006; Fletcher and Zwick, 2007; Taylor et al., 2007). In the famous Prisoner's Dilemma, individuals can either cooperate or defect and payoffs are determined by the strategies of the interacting partners. This problem has been intensely studied with several extensions, such as multiple interaction partners, spatial structure, and repeated interactions (Boyd and Lorberbaum, 1987; Farrell and Ware, 1989; Nowak and

May, 1992; Hauert and Schuster, 1997; Nowak, 2006). A major role in the game theoretic approaches is attributed to reciprocity (e.g., Trivers, 1971), which is captured by strategies such as "tit-for-tat" (mirror the behavior of the opponent in the previous round) or "win-stay, lose-shift" (change strategy when performing poorly) in the repeated Prisoner's dilemma (Nowak, 2006). A review of conditions favoring cooperation under direct, indirect, and network reciprocity as well as under kin and group selection is provided by Nowak (2006). Examples of reciprocity in nature are given by Milinski (1987), McGregor (2005), and Bshary and Grutter (2006). The Prisoner's dilemma game is used by Turner and Chao (1999) to explain the evolution of selfishness in an RNA virus. Furthermore, the evolution of cooperation is investigated in public goods games (see Archetti and Scheuring, 2012 for a review). Here, cheating takes the form of exploiting a public good, which can lead to a phenomenon known as the "tragedy of the commons" (Hardin, 1968). In the original description of Hardin (1968), a pasture is described that can be used freely by several farmers. For any single farmer it is advantageous to bring as many cattle as possible. Thus, the common ground can get exploited and thereby be rendered useless for all farmers. In general, when selfish behavior leads to the extinction of a population, this is an instance of "evolutionary suicide" (Parvinen, 2010).

1.1.3.5 Recognition

Apart from the theories discussed above, the occurrence of altruism can be explained if altruistic behavior is linked to some trait that can be recognized by others (Hamilton, 1964b, pp. 24–25). Then, altruists can target their helping behavior to other carriers of that trait, which diminishes the advantage of non-altruistic free-riders. Dawkins (1976) elaborates on this idea with the example of a hypothetical green beard. Hence, this concept became known as the "greenbeard effect" (see Gardner and West, 2010, for an overview). There still remains the possibility of cheating, if individuals can obtain the green beard without behaving altruistically, which at first led to the conclusion that the practical relevance of the greenbeard effect is low (Grafen, 1984; West et al., 2007a). Nevertheless, there are several examples—predominantly in microorganisms—of greenbeards in nature. These examples are reviewed by Gardner and West (2010), who conclude that the importance of greenbeard effects may have previously been underestimated.

1.2 Altruistic defense

1.2.1 Parasitism

Parasitism is a wide-spread phenomenon in nature and coevolutionary processes are major factors in the shaping of biodiversity (Thompson, 1999). The impact of parasites on host species can be massive and the coevolution of host and parasite can be interpreted as an “arms race” (Dawkins and Krebs, 1979), capturing the fact that higher parasite virulence can be countered by improved host defense. Such reciprocal evolution can put high selective pressure on both hosts and parasites. Similarly, the “red queen” hypothesis (Van Valen, 1973, 1974) describes how evolution of hosts and parasites may cancel each other out, such that the system can remain stable over a long period of time.

A famous approach to describe host-parasite interactions is to model population size changes with differential equations, known as Lotka-Volterra equations (Lotka, 1920; Volterra, 1926). Denoting the population sizes of hosts and parasites by h and p , respectively, and choosing appropriate model parameters $\alpha, \beta, \gamma, \delta \in \mathbb{R}^+$, the changes over time are then given by

$$\begin{aligned}\frac{dh}{dt} &= \alpha h - \beta hp, \\ \frac{dp}{dt} &= -\gamma p + \delta hp.\end{aligned}\tag{1.5}$$

Equation (1.5) is the simplest form of Lotka-Volterra dynamics and it captures the fluctuations of host and parasite population sizes, until, depending on the choice of parameters, a stable equilibrium is reached. There exist various extensions, such as adding multiple populations, intra-species competition, spatial structure, or stochasticity.

Spatial structure can have a strong impact on the dynamics of host-parasite systems (Débarre et al., 2012) and various models of non-panmictic populations have been studied in recent years with computer simulations (e.g., Haraguchi and Sasaki, 2000; Rauch et al., 2002, 2003; Best et al., 2011).

A special case is that of “social parasites”. These are social insects that parasitize and exploit whole societies. Typically, social parasites and their hosts are closely related species. This observation is termed “Emery’s rule” (Emery, 1909). By exploiting the social system of another species, the cost of parental care can be reduced or completely avoided. Among the approximately 13 000 known species of ants, there are roughly 200 that are socially parasitic, but note that the number of extant species may be much higher (Buschinger, 1986; Brandt et al., 2005a; Ward, 2014). A review of the different types of social parasites in ants is given by Buschinger (1986). One example of social parasitism is that of slavemaking

in ants, described in Section 1.2.3. The high level of relatedness between social parasites and their hosts makes close coevolutionary dynamics and a stepwise arms race especially likely (Foitzik et al., 2003). Another special case apart from social parasites is that of avian brood parasites such as cuckoos (Davies et al., 1989) and cowbirds (Ortega, 1998) who lay their eggs in nests of other species. They trick the host birds into caring for their brood, by mimicking visual and acoustic cues, such as color patterns for eggs (Davies and Brooke, 1989) or calls of the young (Langmore et al., 2003).

1.2.2 Altruistic defense

Due to the large impact of parasites, the host species can greatly benefit from improved defenses. Depending on the system that is studied, defense mechanisms can take numerous forms and it can even be associated to some cost. Boots and Bowers (2004) investigate evolution of costly immunity, while costly resistance is considered by Boots and Haraguchi (1999). Duncan et al. (2011) give an overview of studies that show cost of resistance against parasites in various species. Furthermore, Duncan et al. (2011) experimentally demonstrate an association of resistance against a parasite with lower reproductive output in *Paramecium Caudatum*. After contact to the parasite is removed, the selection pressure to maintain the costly defense is eliminated and productivity slowly increases over time. A review of various modeling approaches of different defense mechanism with an associated cost is given in Boots et al. (2009). An overall conclusion is that costly defense behavior can be worthwhile, if selection pressure from parasites is strong enough. Hence, altruistic defense can play an important role in the coevolutionary process of hosts and parasites. Especially in the context of a strong population structure, resistance is beneficial for the hosts (Best et al., 2011). Determining conditions for the evolution of altruistic defense traits can be difficult with existing methodology. Benefits of altruistic defense are not provided immediately, but are mediated by the parasite population. Thus, there may be a time lag between an act of defense and its benefit for the host population, and an entangling of effects of the different actions can be challenging.

There are numerous examples of altruism taking the form of defense against a parasite or predator. When a pathogen infects bacteria, these bacteria can commit suicide, which prevents secondary infections and thereby protects the host population (Labrie et al., 2010). Such a suicidal defense mechanism against infection can be selectively favored in a spatially structured population, as shown experimentally and with simulations by Fukuyo et al. (2012). Suicidal defense in bacteria is also investigated by Berngruber et al. (2013) with a theoretical model and experimental verification of the results. The model is—similar to other studies

on host-parasite interactions—based on deterministic Lotka-Volterra equations (Boots and Haraguchi, 1999; Boots and Bowers, 2004; Miller et al., 2007; Best et al., 2009). Typically, these models describe the numbers of host individuals in different states. Individuals change their states from susceptible to infected, sometimes also to recovered or back to susceptible. Hence, such models are called SI, SIR, SIS, or SIRS (see Boots et al., 2009 for further examples). Berngruber et al. (2013) state that, contrary to other forms of altruism, altruistic defense depends on parasite presence, which implies that benefits of altruistic acts are not constant. In particular, this leads to the conclusion that altruistic defense only evolves at intermediate levels of mixing that simultaneously allow for a spread of parasites and for a population structure that is sufficiently strong for altruism to be selectively favored. Similar to the case of bacteria, social insects show various self-destructive defense behaviors against attackers (reviewed by Shorter and Rueppell, 2012). Ground squirrels can make alarm calls when they detect an approaching predator (Sherman, 1977). The caller may put himself at danger of being noticed first by the predator, but nearby squirrels are warned and get a chance to seek shelter. This behavior is mainly observed when the group consists of close kin (Sherman, 1977; Dunford, 1977). Recently, slave rebellion was discovered as a defense mechanism of ants against their slavemakers (Achenbach and Foitzik, 2009), which can be interpreted as another instance of an altruistic defense trait (Pamminger et al., 2014).

1.2.3 The slavemaking ant *Protomognathus americanus* and its hosts

The slavemaking ant *Protomognathus americanus* and its closely related three host species *Temnothorax longispinosus*, *T. curvispinosus*, and *T. ambiguus* occur in the northeastern part of North America. They live in mixed deciduous forests and build their nests in acorns, small hollow sticks, or hickory nuts on the ground (Achenbach and Foitzik, 2009). New colonies of *P. americanus* are founded by a mated queen that invades a host nest and kills or drives out all queens and adult workers that are present in the nest (Wesson, 1939). This happens during mid to late summer, when there are worker pupae in the host nest (Foitzik and Herbers, 2001b).

There is a critical time of chemical learning for young host workers shortly after eclosion, during which they can strongly be influenced by allospecific odors and learn the recognition cues of their parasites (Hare and Alloway, 1987; Achenbach et al., 2010). Henceforth, the enslaved *Temnothorax* workers carry on to perform their tasks of foraging and brood care in the slavemaker nest.

Further manipulation of host ants by their parasites is achieved by physical aggression (Heinze et al., 1994) and via the use of glandular secretions as chemical weapons to induce

fight between host individuals (Foitzik et al., 2003).

Typically, *P. americanus* nests contain a single queen and only a few slavemaker workers (Herbers and Stuart, 1998), while nests of *T. longispinosus* can contain zero, one, or multiple queens (Alloway et al., 1982). Colonies of *P. americanus* and their hosts are polydomous, i.e., they can span multiple nest sites that exchange queens, workers, and brood (Alloway et al., 1982). *P. americanus* is an obligate social parasite, meaning that it needs its host species to survive. Workers of *P. americanus* are not able to maintain the nest, forage, care for the brood, or even feed themselves (Wesson, 1939; Stuart and Alloway, 1985). Instead, their task is to replenish the slave count of the colony (Buschinger, 1986). In order to do so, they scout for nests of the host species in the vicinity. When a host nest is found, the slavemaker worker then gathers more workers and slaves from the colony and together they raid the host nest. They kill or drive out all queens and adult workers that are present in the nest and then proceed to steal the pupae and transport them to the slavemaker colony. The stolen brood will then join the slave force. In a slavemaker population studied by Foitzik and Herbers (2001b), a mature nest performs an estimated number of six raids per year. There is a high frequency of two thirds of males being produced by workers of *P. americanus* (Foitzik and Herbers, 2001a), while worker reproduction is less important in the host species (Herbers and Cunningham, 1983). The impact of slavemakers on the host species can be massive, with the chance of a single colony being raided in any given year going up to over 50% (Foitzik and Herbers, 2001b). Furthermore, Foitzik and Herbers (2001a) find no local survivors of raided host nests and they conclude that raids completely destroy the targeted host nest or force all survivors to move far away. In either case, slavemaker presence can dramatically alter the fate of a host colony. A study by Foitzik et al. (2001) finds local variation regarding severity of raids between populations of *P. americanus* as well as local variation in effectiveness of defense between *T. longispinosus* populations. This is consistent with the idea of a geographic mosaic of coevolution (Thompson, 1999), which can be caused by population structure. Across the geographic range, there are coevolutionary hot and cold spots, referring to the strength of selective pressure due to the host-parasite interactions. Such a pattern has been identified numerous times for interactions between slavemaking ants and their hosts (Blatrix and Herbers, 2003; Brandt and Foitzik, 2004; Fischer and Foitzik, 2004). One contributor to this geographic mosaic of coevolution may be the variation in slavemaker presence of different species, as shown by Johnson and Herbers (2006) for interactions of *P. americanus* and *T. duloticus*.

1.2.4 Slave rebellion

Gladstone (1981) argues that slave rebellion would never evolve in ants. He bases this conclusion on the fact that slaves never reproduce and on the possibility that a refusal to work could induce the slavemakers to go on more raids. These two points mean that rebellion would not give a direct fitness advantage to slaves and could even be harmful to nearby host nests.

Nevertheless, Achenbach and Foitzik (2009) give evidence for rebellion among slaves of *T. longispinosus*. Slaves attack and kill pupae of their slavemaker *P. americanus* instead of caring for them. This behavior is observed across different geographic locations, with strong variation in its intensity (Pamminger et al., 2013).

The cuticular hydrocarbon profiles of pupae of *P. americanus* and their *Temnothorax* hosts are highly divergent (Achenbach et al., 2010). Hydrocarbons are used by insects for many different functions (see Howard and Blomquist, 1982, 2005, for an overview). One of these functions is recognition of species and nest membership (Lahav et al., 1999). The highly divergent profiles can thus be used by enslaved host workers to identify and kill pupae of their parasite, instead of raising them. *P. americanus* cuticular hydrocarbon profiles resemble those of their local hosts more closely than those of distant ones. This indicates that the rebellion behavior of enslaved host workers evolved only recently and that now there is strong selective pressure on parasites to adapt by mimicking cuticular hydrocarbon profiles of local host populations (Achenbach et al., 2010). Hence, in this case, the hosts are apparently one step ahead in the coevolutionary arms race.

Pamminger et al. (2014) show that the rebellion behavior can be an instance of altruistic defense. Having the defense trait may be associated with a cost to the individual for developing the necessary structures. Furthermore, carriers of the rebel allele may mistake conspecific offspring in free-living nests for enemies and thus they might be prone to accidental killings in their mother nests (Metzler et al., 2015). Such erroneous killings of own offspring occur in hosts of avian brood parasites (Davies et al., 1996). This is caused by the strong selective pressure induced by their parasites to accept only a very limited range of phenotypes of eggs in their nests.

Viewing a whole nest as a “superorganism” (Hölldobler and Wilson, 2009) permits the idea of slave rebellion being an altruistic trait. Free-living nests that contain rebels pay the cost of a reduced growth rate. They benefit host nests in the area by reducing slavemaker pressure, once their workers get enslaved and kill slavemaker offspring. Similarly, the queens themselves can be seen as the agents in this system. A queen that bears the rebel allele has a nest with rebel workers and the above reasoning applies. Hence, even though the slaves

themselves do not get a direct fitness advantage from rebellion, the trait could still evolve via indirect fitness benefits, even if it is costly (Pamminger et al., 2014).

1.3 Goals and outline

The overall goal of this dissertation is to better understand the evolution of altruism in the form of defense against a parasite or predator. This is achieved in three parts.

The first part consists of the mathematical analysis of a model for the evolution of an allele for altruistic defense in a population that is structured in demes. In this part, the goal is to derive an analytical result on the success of an altruistic defense allele. In Section 2.1, the model that forms the basis of the analysis is introduced and the ideas and techniques of the mathematical analysis are described. All technical details and proofs, as well as intermediary results, are provided in Chapter 4. The starting point is an individual-based model of altruists, cheaters, and parasites, which is introduced in Section 2.1.1. Interactions of host and parasite (or prey and predator), as well as the population structure, make it difficult to determine the success of altruists. Thus, an important first goal is to obtain an expression that describes the frequency of altruists. To that end, the diffusion approximation of the individual-based model is derived in Section 4.1.2. The resulting system of diffusions is shown in Section 4.1.3. In Section 4.1.4, all intermediary results and all proofs in the derivation of the process X , as given by equation (3.2), are shown. This process describes the relative frequency of altruists in each deme in the limit of large populations under the assumption of a complete separation of ecological and evolutionary time scales. Further simplification is achieved in Section 4.1.5 by assuming uniform migration across all demes and taking the limit of many demes. This leads to a single one-dimensional process describing the average altruist frequency. Section 4.1.6 then explores the longterm behavior of this process. A condition for the success of an altruistic defense allele is derived that assumes positive frequency of the allele in an infinite number of demes with infinite population sizes. Finally, the invasion of an altruistic defense allele in a population of non-altruists is analyzed in Section 4.1.7 and a condition for possible survival of the allele after arising at low frequency in a single deme is derived. The ideas and techniques of the steps described above are presented in a less technical manner in Sections 2.1.2 to 2.1.7. All key results are provided in Section 3.1.

The goal of the second part is to evaluate the results from part one in relaxed settings. Deriving the result in the first part includes taking the limits of infinite population sizes, infinitely many demes, and an infinite time horizon. Hence, the result can—at best—only approximate the situation of any natural population. In fact, it is possible that taking these limits qualitatively changes the outcome of the system compared to the finite setting. The

amount of deviation from the theoretical prediction is checked in the second part with computer simulations of two different models that are described in Sections 2.2.1.1 and 2.2.2.1, respectively. Both models assume finitely many demes and (of course) a finite time horizon. First, the individual-based model with finite population sizes that forms the starting point of part one is simulated. Second, the process X , describing the altruist frequency under infinite population sizes and a complete separation of time scales is approximated in a discretization scheme. This allows for an investigation of the differences between finite and infinite population sizes as well as the effect of a separation of time scales. To better understand how the success of altruists can be influenced and check the robustness of the results, some extensions are added to both models and various parameter settings are explored. Details of the simulation studies are provided in Sections 2.2.1.2 and 2.2.2.2, respectively. Results are presented in Section 3.2.

The third part examines a specific altruistic defense trait found in nature—slave rebellion in ants. Here, the goal is to understand how the situation in a natural population may influence the results from the theoretical prediction and to find mechanisms and parameters that are relevant for the success of this specific altruistic defense trait. In order to obtain a model that closely resembles the situation in the field, many details regarding the ecology of the system of slavemakers and their host are incorporated. A patch of forest with individual host and slavemaker ants, organized in nests, is modeled. All details of the model are given in Section 2.3.1. The different simulation studies are described in Section 2.3.2. First, in Section 2.3.2.1, suitable parameter sets are determined using field data. With these parameter sets, several simulation series are performed to find conditions that allow a spread of rebellion, under variable cost settings (Sections 2.3.2.2 and 2.3.2.5) and fixed costs (Sections 2.3.2.3 and 2.3.2.6). To verify that rebellion can only spread in the presence of slavemakers and is not a purely random phenomenon, additional simulations (described in Section 2.3.2.4) are performed. In Section 2.3.2.7, a metapopulation with several patches of forest that are connected via migration is studied to check how this additional level of population structure influences the evolution of rebellion. Results are shown in Section 3.3.

In Chapter 5, the findings of the three parts are discussed.

Chapter 2

Models and methods

2.1 Mathematical analysis and asymptotics

In this section, a stochastic model is introduced and the ideas and techniques of the mathematical analysis, which is performed in Chapter 4, are described.

2.1.1 Individual-based model

In this section, an individual-based model for the evolution of an altruistic defense trait is defined. Assume there is a parasite population and a host population with two types of host individuals—altruists and cheaters—and that these populations are structured in demes that exchange migrants.

The model describes the numbers of altruists, cheaters, and parasites. For each type, birth, death, and migration rates are specified. These rates incorporate randomness in reproduction, competition within species (to account for limited space and resources per deme), and interactions between hosts and parasites. Hosts suffer from parasite presence, while parasites need hosts to survive and grow. Altruists provide the benefit of locally reducing parasite growth and pay the cost of an increased death rate compared to cheaters. Starting from an initial configuration of altruists, cheaters, and parasites in each deme, the populations evolve according to the transition rates given in equation (2.1) below. These rates refer to exponentially distributed waiting times of the transition events. One transition event consists of a single reproduction, death, or migration of one individual. All parameters of the model are shown in Table 2.1.

Denote the set of all demes by \mathcal{D} , which is assumed to be finite (e.g., $\mathcal{D} = \{1, \dots, D\}$, for some $D \in \mathbb{N}$) or countable (e.g., $\mathcal{D} = \mathbb{N}$). Assume that the numbers of altruists, cheaters, and parasites are $a_k, c_k, p_k \in \mathbb{N}_0$ (where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$), in each deme $k \in \mathcal{D}$, respectively.

Parameter	Description	Domain
λ	host growth rate	\mathbb{R}^+
K	host carrying capacity	\mathbb{R}^+
δ	host death per parasite	\mathbb{R}^+
α	additional death rate of altruists	\mathbb{R}^+
ν	parasite death rate	\mathbb{R}^+
γ	parasite competition	\mathbb{R}^+
η	parasite growth per host	\mathbb{R}^+
ρ	parasite death per altruist	$(0, \eta)$
κ_H	host migration rate	\mathbb{R}^+
κ_P	parasite migration rate	\mathbb{R}^+
β_H	randomness in host reproduction	\mathbb{R}^+
β_P	randomness in parasite reproduction	\mathbb{R}^+
N	population size scaling	\mathbb{N}
D	number of demes	$\mathbb{N} \cup \{\infty\}$
T	time horizon	\mathbb{R}^+
m	migration matrix	$[0, \infty)^{D \times D}$
A_0	initial altruists per deme (units of N)	\mathbb{N}
C_0	initial cheaters per deme (units of N)	\mathbb{N}
P_0	initial parasites per deme (units of N)	\mathbb{N}

Table 2.1: Parameters of individual-based model.

Then the transition rates are given by

$$\begin{aligned}
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k + \mathbb{1}_{k=i}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \left(\beta_H + \lambda \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k - \mathbb{1}_{k=i}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \left(\beta_H + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} + \frac{\alpha}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k + \mathbb{1}_{k=i}, p_k)_{k \in \mathcal{D}}: & c_i \left(\beta_H + \lambda \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k - \mathbb{1}_{k=i}, p_k)_{k \in \mathcal{D}}: & c_i \left(\beta_H + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k + \mathbb{1}_{k=i})_{k \in \mathcal{D}}: & p_i \left(\beta_P + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k - \mathbb{1}_{k=i})_{k \in \mathcal{D}}: & p_i \left(\beta_P + \nu + \gamma \frac{p_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \frac{\kappa_H}{N} m(j, i) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}, p_k)_{k \in \mathcal{D}}: & c_i \frac{\kappa_H}{N} m(j, i) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j})_{k \in \mathcal{D}}: & p_i \frac{\kappa_P}{N} m(j, i),
 \end{aligned} \tag{2.1}$$

where $\mathbb{1}_{k=i} = 1$, if $k = i$ and $\mathbb{1}_{k=i} = 0$, otherwise. Note that for all $i, j \in \mathcal{D}$, $m(j, i)$ describes migration from deme i to deme j . It is assumed for every $i \in \mathcal{D}$ that $\sum_{k \in \mathcal{D}} m(k, i) =$

$\sum_{k \in \mathcal{D}} m(i, k) = 1$. In this model, the transition rates at any point in time only depend on the current state of the process. Such a process for which only the current state is relevant to determine the next transition is called a “Markov process” (e.g., Klenke, 2008, who also provides a general introduction to probability theory). Because the past history is not taken into consideration, Markov processes are also said to be “memoryless”.

The model will be analyzed analytically in the remainder of Section 2.1. All technical details are provided in Chapter 4. A simulation approach of the model is described in Section 2.2.1.

2.1.2 Diffusion approximation

In this section, the approach of Section 4.1.2 is described. The model presented in Section 2.1.1 is hard to analyze. Hence, the first step is to transform the discrete model to a continuous “diffusion process” (e.g., Klenke, 2008). This process takes the form of stochastic differential equations, describing changes in population sizes continuously in time. The changes consist of deterministic components, to account for the trend of how the population sizes change due to ecological and evolutionary forces, and random components, to describe the uncertainty in reproduction and death events (“genetic drift”). The transformation of the model is achieved by a “diffusion approximation” (Ethier and Kurtz, 1986). Examples of diffusion approximations for the Wright-Fisher model and the Moran model are given in Chapter 7 of Durrett (2008).

For the model introduced in the previous section, individuals are added or removed at discrete points in time and the process makes jumps of size one. By dividing the numbers of individuals by N , the jumps are reduced to size $1/N$ and for large N , a continuous process can be approximated. Thus, instead of jumps at discrete points in time, small changes occur continuously for the resulting diffusion, which approximates the original Markov process. A representation of the diffusion limit is computed via equation (4.2).

2.1.3 Diffusion model

In Section 4.1.3, the diffusion model for altruists, cheaters, and parasites, which is obtained by the diffusion approximation, is formally described. Equation (4.9) gives the numbers of altruists, cheaters, and parasites, at time t , in deme i , counted in units of N individuals—e.g., $A_t^N(i) = 2$ means that in deme i there are $2N$ altruists at time t . Note that these processes are continuous approximations of the birth-death process of Section 2.1.1, though, and thereby (most of the time) are not integer-valued, even after multiplication by N . An introduction

to diffusions and stochastic differential equations can be found in Klenke (2008), Karlin and Taylor (1981), or Øksendal (1998). In fact, equation (4.9) belongs to the class of Lotka-Volterra equations (Lotka, 1920; Volterra, 1926), which describe host-parasite or predator-prey dynamics. For these types of equations, the interactions of host and parasite (or predator and prey) will lead to fluctuations over time until an equilibrium is reached. The equilibrium values can be obtained by setting the amount of change in both types to zero. However, in the present scenario, deriving these equilibrium values is complicated, as spatial interactions with random components are considered.

2.1.4 Convergence of the relative frequency of altruists

In Section 4.1.4, a process X , describing the relative frequency of altruists among hosts, is approximated from the diffusion model for large population sizes. This simplifies the further analysis, because then only X needs to be understood, instead of simultaneously observing A , C , and P .

First, an equivalent representation of the processes A , C , and P is derived in Lemma 4.1.2, such that the numbers of altruists, cheaters, and parasites are described by H , F , and P , where $H = A + C$ describes the total number of host individuals and $F = A/(A+C)$ describes the relative frequency of altruists among all hosts. Thus, the goal is to determine the dynamics of F , which still depends on H and P .

In Lemma 4.1.3, it is shown that (under certain assumptions) hosts will never go extinct in any of the demes. In fact, this result is shown to hold with probability one, which for practical purposes means that it can be assumed to always hold (see “almost surely” and “null set” in Klenke, 2008, Chapter 1). The result is obtained by comparing H with a similar process, which is easier to analyze and which is shown to provide a lower bound for H . In Lemma 4.1.4, the analogous result is stated for P .

In population genetics, a typical assumption is that of weak selection, weak migration, and weak genetic drift. This is achieved by scaling the corresponding terms with the inverse of the population size and considering large populations. Consequently, it takes a long time to observe an effect of these evolutionary forces. In contrast, the ecological forces of reproduction, death, host-parasite interactions and intra-species competition are unscaled in the model and thus, act much faster. Hence, the equilibrium values for H and P —which can be derived analogously to the classical Lotka-Volterra equations (Lotka, 1920; Volterra, 1926), but which here depend on F —are reached before the evolutionary forces show any effect. Meanwhile, F is not influenced by the ecological interactions in large populations. Thereby, a complete “separation of time scales” is achieved and it takes a long time until a change in

F occurs. Once this change becomes noticeable, H and P are already at their equilibrium values. By speeding up the processes appropriately, ongoing changes occur in F while H and P stay at their equilibrium values at all times. This intuition is made rigorous in Theorem 3.1.1, which rests on a so-called “Lyapunov function”. This is a non-negative function that decreases along trajectories of the considered differential equation and is zero at the equilibrium, while being positive at every other point (e.g., Wirsching, 2006). In Theorem 4.1.9, this Lyapunov function is derived analogous to the standard Lotka-Volterra case. The proof of Theorem 4.1.9 relies on the boundedness of the expected values of several terms of H and P , which is shown in Lemma 4.1.6, Lemma 4.1.7, and Lemma 4.1.8. In these lemmas, the interactions between hosts and parasites—according to the Lotka-Volterra fluctuations—have to be accounted for. If H gets larger, then P grows faster, which in turn leads to a decrease in H . Thus, in order to show boundedness of the relevant expected values, the processes have to be considered simultaneously by combining them in a suitable way. An important tool in proving these results is Itô’s lemma (e.g., Klenke, 2008), which allows to derive the dynamics for functions of H and P . Finally, the proof of Theorem 3.1.2, which rests on the above-mentioned results, is provided. It formalizes the idea that the processes describing the numbers of hosts and parasites can be substituted by their equilibrium values to obtain a description of the altruist frequency on the evolutionary time scale.

The results in Section 4.1.4 are shown for countable sets of demes, which are infinitely large. An example of a countable set is \mathbb{N} , the set of all natural numbers. In this setting, simply summing the number of individuals over all demes would (typically) lead to infinite amounts and thus not allow for comparing different configurations or obtaining a useful notion of boundedness. Thus, it is necessary to introduce a way of assigning sizes or weights to configurations of individuals in these infinitely many demes. Thus, the σ -norm ($\|\cdot\|_\sigma$) is introduced, which assigns weights to all demes in a way that typical configurations (e.g., at most M individuals per deme, for some $M \in \mathbb{N}$) have finite sizes. The results apply to finite sets of demes as well, where the σ -norm could simply be replaced by the sum over all demes.

2.1.5 McKean-Vlasov limit

In Section 4.1.5 the process X , as derived in Section 4.1.4, is simplified further. For this, it is assumed that there are D demes, which are connected by uniform migration. The idea here is that the terms describing immigration events in the case of uniform migration are of the form $\frac{1}{D} \sum_{j=1}^D \psi(X_t^D(j))$, for some function ψ . This means that these immigration terms are averaged over all demes. In the case of many demes, this expression will approximate

the expected value, $\mathbb{E}[\psi(X_t^D(i))]$. Thus, the average over all demes can be replaced by the expectation in the limit of many demes, which greatly reduces the complexity of the process (from D dimensions to only one dimension). The result is first presented in a general setting in Proposition 4.1.13. Applicability to the present scenario of the altruistic defense allele is then verified in Lemma 4.1.14.

2.1.6 Long-term behavior of the average altruist frequency

In Section 4.1.6, the long-term behavior of the process Z describing the average altruist frequency in the case of many demes with uniform migration is investigated. This process contains an expectation term, stemming from the average over many demes, as described in the previous section. The idea is to use the insight that if the process is at equilibrium, then this expected value stays constant over time (Lemma 4.1.15). Hence, Z is compared to another process, \hat{Z} , for which the expectation term is replaced by a constant, which makes this second process easier to analyze. If both processes start at the same value, then it is shown that the expectation term in Z gets larger or smaller than the constant term in \hat{Z} , depending on α and β . This is true for any constant term and it is concluded that Z goes to extinction or fixation, depending on α and β . In the critical case $\alpha = \beta$, there is an equilibrium of altruists and cheaters and the distribution of the types is specified in Theorem 3.1.3.

2.1.7 Invasion of an altruistic defense allele

Thus far, the assumption is that both altruists and cheaters are present in all demes. This does not consider the case of a recently arisen altruistic allele. Hence, in Section 4.1.7 it is assumed that the altruistic allele is present only in a single deme and in low frequency, possibly due to a recent mutation event. The aim is to determine the possibility of invasion in a structured population of non-altruistic individuals. Large populations are assumed and the corresponding process for the altruist frequency (as derived in Section 4.1.4) is analyzed under uniform migration. If, in the beginning, the altruist frequency is positive only in a single deme and there is a large number of demes, then it is very rare that altruists migrate to a deme where other altruists are already present. Hence, this case can be neglected (until altruists have established in a sufficient number of demes) and the immigration term can be set to zero in the process describing the altruist frequency in any deme with altruists. In Proposition 4.1.17, this simplification is used to derive the condition for certain extinction of altruists.

2.2 Simulations on the evolution of altruistic defense

2.2.1 Small populations

2.2.1.1 Simulation model

The evolution of altruists, cheaters, and parasites is modeled on a set of D demes. The simulation scheme is outlined in Box 2.1 and all model parameters are described in Table 2.2. First, the initial sizes of altruists, cheaters, and parasite in each deme are set to A_0N ,

Box 2.1: Simulation scheme for simulations with finite population sizes.

```

Set  $a_i := A_0N$ ,  $c_i := C_0N$ ,  $p_i := P_0N$ , for all  $i \in \{1, \dots, D\}$ 
Set  $t := 0$ 
while  $t < T$  &  $\sum_{i=1}^D a_i > 0$  &  $\sum_{i=1}^D c_i > 0$  do
    Set  $t := t + \frac{1}{r}$ , where  $r$  is the sum of all rates
    Pick random transition with probabilities
    proportional to rates (see Box 2.2)
    Update population sizes according to the chosen
    transition event
    Update transition rates with new population sizes
end while

```

C_0N , and P_0N , respectively. In each simulation step, a single transition is performed, which can be a birth, death, or migration event for one of the three types in any deme. First, the inverse of the sum of all transition rates is added to the time that has passed. This corresponds to the expected value of the time until the next transition occurs, which is used instead of the true value (which is exponentially distributed with a rate of the sum of all transition rates) for computational efficiency. The transition is randomly chosen with probabilities proportional to the transition rates, which are given in Box 2.2. After the numbers of individuals are updated with respect to the randomly chosen transition event, the transition rates are adjusted accordingly. If the time that has passed exceeds the specified time horizon or if there is a fixation of altruists or cheaters, then the simulation run terminates. Otherwise, the next transition step is performed. The process of finding the next transition is split in two steps. These steps involve, first, identifying the corresponding deme which the transition occurs in and second, determining one transition among all possibilities in that deme. For computational efficiency, a binary tree is used in the first step to find the deme for the next transition. The tree has one leaf per deme and the sum of all rates of each deme is stored in the corresponding leaf. Any node holds the sum of all values of the leaves that descend

Box 2.2: Transition rates when the numbers of altruists, cheaters, and parasites are a_k , c_k , and p_k , respectively, in all demes $k \in \mathcal{D}$, with $\mathcal{D} := \{1, \dots, D\}$. For all $i \in \mathcal{D}$, the rates are given by

$$\begin{aligned}
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k + \mathbb{1}_{k=i}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \left(\beta_H + \lambda \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k - \mathbb{1}_{k=i}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \left(\beta_H + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} + \frac{\alpha}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k + \mathbb{1}_{k=i}, p_k)_{k \in \mathcal{D}}: & c_i \left(\beta_H + \lambda \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k - \mathbb{1}_{k=i}, p_k)_{k \in \mathcal{D}}: & c_i \left(\beta_H + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k + \mathbb{1}_{k=i})_{k \in \mathcal{D}}: & p_i \left(\beta_P + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k - \mathbb{1}_{k=i})_{k \in \mathcal{D}}: & p_i \left(\beta_P + \nu + \gamma \frac{p_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \frac{\kappa_H}{N} \frac{1}{D} \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}, p_k)_{k \in \mathcal{D}}: & c_i \frac{\kappa_H}{N} \frac{1}{D} \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j})_{k \in \mathcal{D}}: & p_i \frac{\kappa_P}{N} \frac{1}{D}
 \end{aligned}$$

from that node. This means that the root stores the total sum of all rates. Its left child holds the sum of the rates for the first half of all demes and its right child holds the corresponding sum for the second half of all demes. If a node has an odd number of descending leaves, the left child has one descendant more. The deme for the next transition is then found in the following way. A random real number r is drawn uniformly between zero and the sum of all transition rates. The active node is set to be the root. If r is smaller than the value of the left child of the active node, then the active node is set to be the left child. Otherwise, the value of the left child is subtracted from r and the active node is set to be the right child. This procedure is repeated until the active node is a leaf and the deme for the next transition is the one corresponding to that leaf. In the second step, the transition is determined. If r is less than the sum of all rates of altruists, i.e., birth, death, and migration rate, then the transition involves the altruists. In that case, if r is less than the altruist birth rate, the number of altruists in the corresponding deme is increased by one. Otherwise, the altruist birth rate is subtracted from r . Then, if r is less than the altruist death rate, the number of altruists in the corresponding deme is decreased by one. Otherwise, a migration occurs and the number of altruists in the corresponding deme is decreased by one and the number of altruists in a random deme, drawn uniformly among all demes, is increased by one. Note that the target deme of the migrant may also be the current deme. In that case nothing happens. If the

transition does not involve altruists, the procedure is repeated for cheaters and parasites until the transition is found. An outline of how the next transition is found is given in Box 2.3.

2.2.1.2 Simulation study

Simulations with different parameter settings were performed in order to check if the prediction from the theoretical analysis holds in the case of finite populations in finitely many demes.

In the finite setting, randomness is inherent to the simulation model, even without the added randomness introduced by β_H and β_P . This may mean that a disfavored type that would go extinct in the infinite model may prevail in the finite setting due to chance. To check this effect, *simulation series* α was performed with different values of the altruist disadvantage α and different values of the effect of altruistic defense, ρ , which led to different values of the “benefit”, $b = \rho \frac{\delta}{\delta\nu + \lambda\gamma}$ (i.e., $b = \beta/\beta_H$).

To further understand how this randomness in the finite model affects the outcome, *simulation series* β_H was performed for different values of β_H .

In the setting of $N < \infty$, the time scales of evolutionary and ecological forces are not separated. Thus, different migration rates may play a more pronounced role. A low value of κ_H may mean that effects within demes dominate the dynamics, while a very large value may remove the effect of population structure. The influence of κ_H was investigated in *simulation series* κ_H .

In *simulation series* a_f , the assumption of having identical conditions in each deme was relaxed. After each year, each deme had a change in the competition rates, independently for hosts and parasites, both with probability of p_f . If such a change occurred, a normally distributed random number r with mean zero and standard deviation $a_f\lambda/KN$ for hosts or $a_f\gamma/N$ for parasites was drawn. The new competition rate was then set to $\lambda/KN + r$ in the case of hosts or $\gamma/N + r$ in the case of parasites. If the resulting value was non-positive, a new random number was drawn (until a positive value was obtained). The effect of these fluctuations was investigated for different values of a_f .

The values for all parameters in the different simulation series are shown in Table 2.2. All data was analyzed using *R* (R Core Team, 2015). Wilson confidence intervals were computed with the function *binconf*, logistic regression was achieved with the function *glm* for a binomial distribution with a *logit* link. Kruskal-Wallis tests were performed to check for significance using the function *kruskal.test* and p-values were adjusted for multiple testing with the Bonferroni-Holm correction via the function *p.adjust*.

Box 2.3: Procedure to determine the next transition event for simulations with finite population sizes.

Draw uniform real number r between zero and sum of all rates

Set active node to be the root of the binary search tree (described in text)

while active node not leaf **do**

 Set r_l to be the value of left child of active node

if $r < r_l$ **then**

 Set active node to be left child of active node

else

 Set $r := r - r_l$

 Set active node to be right child of active node

end if

end while

Next transition occurs in deme corresponding to leaf that is active node (following rates refer to that deme)

if r less than sum of rates of altruists **then**

 Transition among altruists

else

 Decrease r by sum of rates of altruists

if r less than sum of rates of cheaters **then**

 Transition among cheaters

else

 Transition among parasites

end if

end if

Find transition among type (altruist/cheater/parasite):

if r less than birth rate of type **then**

 Add individual of type to deme

else

 Decrease r by birth rate of type

if r less than death rate of type **then**

 Remove individual of type from deme

else

 Remove one individual of type from deme and add one individual of type to random deme

end if

end if

Parameter	Description	α	β_H	κ_H	a_f
ρ	parasite death per altruist	var	var	var	var
λ	host growth rate	1	1	1	1.5
K	host carrying capacity	2	2	2	2
δ	host death per parasite	1	1	1	0.5
α	additional death rate of altruists	var	1	1	1
ν	parasite death rate	0.5	0.5	0.5	0.5
γ	parasite competition	2	2	2	1.5
η	parasite growth per host	1	2	2	4
κ_H	host migration rate	1	1	var	1
κ_P	parasite migration rate	1	1	1	1
β_H	randomness in host reproduction	1	var	1	1
β_P	randomness in parasite reproduction	1	1	1	1
N	population size scaling	500	500	500	500
D	number of demes	50	50	50	50
T	time horizon	2.5E6	2.5E6	2.5E6	2.5E6
A_0	initial altruists per deme (units of N)	1	1	1	1
C_0	initial cheaters per deme (units of N)	1	1	1	1
P_0	initial parasites per deme (units of N)	1	1	1	1
p_f	probability of fluctuating competition	0	0	0	0.02
a_f	amount of competition fluctuation	0	0	0	var

Table 2.2: Parameters of individual-based model with values used in the different simulation series.

2.2.2 Large populations

Theorem 3.1.2 shows that the frequency of altruists in any given deme converges to the dynamics of the process X , as given by equation (3.2), when time is sped up by a factor of N in the limit of an infinitely large population, $N \rightarrow \infty$. In this section, the simulation procedure of approximating the dynamics of the process X is described.

2.2.2.1 Simulation model

The dynamics of the process X , given by equation (3.2), is approximated by the following simulation scheme. At the beginning of the simulation, X_i is set to x_0 in each deme $i \in \{1, \dots, D\}$ and the current time t is set to zero. The discretization is achieved by choosing a small, positive step size dt , and approximating the change of the Brownian motion in each time step by dW_i , for $i \in \{1, \dots, D\}$, which is normally distributed with mean zero and standard deviation \sqrt{dt} . In each simulation step, corresponding to a time interval of length dt , the value of X_i is updated as

$$X_i := X_i + \left[\kappa(a - X_i) \left(\sum_{j=1}^D m(i, j) \frac{a - X_i}{a - X_j} - 1 \right) - \alpha X_i (1 - X_i) \right] dt + \sqrt{\beta(a - X_i)X_i(1 - X_i)} dW_i, \quad (2.2)$$

for all $i \in \{1, \dots, D\}$, where the migration rate $m(i, j)$ from deme j to deme i depends on the migration scheme that is used. Note that β corresponds to $b \cdot \beta_H$ in the setting of Section 4.1.4.1. If $X_i < \varepsilon$, then X_i is set to zero, or if $X_i > 1 - \varepsilon$, then X_i is set to one (see Box 2.5 for details). Then, t is increased by dt and if t is an integer-valued multiple of $T/1000$, then the average value of X across all demes is computed. If that average value is zero or one, or if $t \geq T$, then the simulation run is terminated. An overview of the simulation procedure is shown in Box 2.4.

2.2.2.2 Simulation study

In order to assess how well the results from the theoretical analysis could predict the outcome under finitely many demes, simulations were performed for a fixed value of altruist disadvantage α and different values of the “benefit” term, β . This was done for different values of the number of demes, D , to check how well the setting of infinitely many demes in the theoretical analysis could be approximated. In further analyses, the robustness of the system towards different migration schemes was investigated. The different schemes were uniform migration, nearest-neighbor migration on a circle, migration to the four immediate

Box 2.4: Simulation scheme for simulations with infinite population size.

```

Set  $X_i := x_0$ , for all  $i \in \{1, \dots, D\}$ 
Set  $t := 0$  and  $avg := \frac{1}{D} \sum_{i=1}^D X_i$ 
while  $t < T$  &  $avg \notin \{0, 1\}$  do
    for all  $i \in \{1, \dots, D\}$  do
        Draw  $dW_i \sim \mathcal{N}(0, \sqrt{dt})$ 
        Set  $X_i := X_i + \left[ \kappa(a - X_i) \left( \sum_{j=1}^D m(i, j) \frac{a - X_i}{a - X_j} - 1 \right) - \alpha X_i (1 - X_i) \right] dt$ 
             $+ \sqrt{\beta(a - X_i)X_i(1 - X_i)} dW_i$ 
        if  $X_i < \varepsilon$  then
            Set  $X_i := 0$ 
        else if  $X_i > 1 - \varepsilon$  then
            Set  $X_i := 1$ 
        end if
    end for
    Set  $t := t + dt$ 
    if  $t \bmod \frac{T}{1000} = 0$  then
        Set  $avg := \frac{1}{D} \sum_{i=1}^D X_i$ 
    end if
end while
    
```

Box 2.5: Cutoff in simulations of process X .

The discretization may introduce a bias away from the boundaries, if X_i attains values very close to zero or one. This is due to the fact that near zero or one, steps in direction of the boundaries cannot exceed the boundaries, while steps away from the boundaries are taken in full size. To compensate for that, values below or above the cutoff values of ε or $1 - \varepsilon$, respectively, are set to the corresponding boundary values of zero or one. In *simulation series* ε , a suitable cutoff value was determined. Here, suitable means large enough such that the bias does not affect the outcome and small enough such that immigration can still drive demes away from fixation or extinction. Note, however, that if the process is simulated for a very long time, this effect may be overcome by chance.

Parameter	Description	ϵ	U	N	2D	T
β	random fluctuation (“benefit”)	var	var	var	var	var
α	altruist disadvantage	0.05	0.05	0.05	0.1	0.05
κ	migration rate	0.1	0.1	0.1	1	0.1
a	see Section 4.1.4.1	2	2	2	2	3
D	number of demes	500	var	var	var	var
x_0	initial altruist frequency per deme	0.5	0.5	0.5	0.5	0.5
T	time horizon	2000	2000	2000	2000	2000
dt	time step size of discretization	1E-5	1E-5	1E-5	1E-5	1E-5
ϵ	cutoff value	var	1E-7	1E-7	1E-7	1E-7

Table 2.3: Parameters of diffusion simulations with values for runs with uniform migration (U), nearest-neighbor migration (N), two-dimensional nearest-neighbor migration (2D), and migration along edges of a binary tree (T). Entries marked with “var” vary for the simulation series and the values are provided in the corresponding plots.

neighbors on a two-dimensional lattice, and migration to connected nodes on a binary tree, where the root was connected to its two children, any inner node was connected to its two children and its parent node, and any leaf was connected to its parent node as well as its two neighboring leaves (i.e., the leaves form a circle along which migration occurs). See Box 2.6 for details on the migration patterns. The parameters used in the different settings are shown in Table 2.3.

Box 2.6: Migration patterns for the simulations series with uniform migration (U), nearest-neighbor migration (N), two-dimensional nearest-neighbor migration (2D), and migration along edges of a binary tree (T). Note that $m(i, j)$ refers to migration from deme j to deme i .

$$\begin{aligned}
 \mathbf{U} \quad & m(i, j) = \frac{1}{D}, \text{ for all } i, j \in \{1, \dots, D\} \\
 \mathbf{N} \quad & m(i, j) = \frac{1}{2}, \text{ if } i \text{ next to } j \text{ on circle; } m(i, j) = 0, \text{ otherwise} \\
 \mathbf{2D} \quad & m(i, j) = \frac{1}{4}, \text{ if } i \text{ next to } j \text{ on two-dimensional lattice; } m(i, j) = 0, \text{ otherwise} \\
 \mathbf{T} \quad & m(i, j) = \frac{1}{3}, \text{ if } j > 1 \text{ is parent or child of } i \text{ or if } i \text{ and } j \text{ are neighboring leaves;} \\
 & m(i, j) = \frac{1}{2}, \text{ if } j = 1 \text{ and } i \in \{2, 3\}; m(i, j) = 0, \text{ otherwise}
 \end{aligned}$$

2.3 Evolution of slave rebellion in ants

2.3.1 Simulation model

In this section, the simulation model of Metzler et al. (2015) is described in detail. A brief outline of the simulation scheme is provided in Box 2.7, Box 2.8, and Box 2.9. Model parameters and variables are summarized in Table 2.6, Table 2.4, and Table 2.5.

In each simulation run, the interactions of slavemaker ants with one host species are simulated on a grid that represents a patch in a forest of $100 \text{ m} \times 100 \text{ m}$. The grid is given a torus structure to avoid boundary effects and it consists of 200×200 grid points, each representing $0.5 \text{ m} \times 0.5 \text{ m}$. The simulated time corresponds to T years, with $T = 5000$ in most simulation series.

During the set up of the simulation run, each grid point is assigned a host nest (i.e., a nest of the host species, irrespective of parasite contact) with probability $8750/40\,000$, or a slavemaker nest with probability $875/40\,000$, or no nest with probability $30\,375/40\,000$. Thus, nest densities are similar to those found in the field (Foitzik et al., 2009). Note that field nest densities vary depending on the location, which is accounted for in the simulations, by simulating a long time span of (at least) 5000 years. This allows for a burn-in time, such that the model can adapt to the specified parameters. If a grid point is assigned a host nest, this nest gets a random number of workers, uniformly distributed between zero and h_{\max} . Furthermore, the nest gets a queen who has the rebel allele on each chromosome with probability f_A . Her age is a randomly chosen integer value between zero and a_h . The queen is mated to a single male, who possesses the rebel allele with probability f_A . Both queen and male are assigned microsatellites on six loci according to empirical distributions found in the field. All loci are assumed to be inherited independently, i.e., without linkage, and mutation is not included in the simulation. If a grid point is assigned a slavemaker nest, this nest gets a random number of slavemaker workers and slaves, uniformly distributed between zero and p_{\max} or s_{\max} , respectively. In addition, the nest gets a queen and her age is a uniformly chosen integer value between zero and a_p . This concludes the set up.

One year corresponds to 40 000 time steps. A single time step includes choosing a random grid point and performing nest actions if a host or slavemaker nest inhibits the chosen point. This happens in each simulation step with probability $\eta/(\eta+q_h+q_p)$, where q_h and q_p denote the numbers of host and slavemaker queens, respectively, currently on their nuptial flights in the area and η determines the rate at which new nests are founded. Because there are 40 000 grid points, each one of them is picked once per year, on average, to perform its actions.

On the other hand, with probability $q_h/(\eta+q_h+q_p)$, a mated host queen is chosen from the pool of queens currently on their nuptial flights, and she attempts to found a new nest on the grid. To that end, a random grid point is chosen. If it is empty, the queen founds a nest there. Otherwise the procedure is repeated up to a total of n_h times. If no empty grid point is drawn, then no new nest will be founded. Independent of the success of the foundation of a new nest, a host queen and a host male are removed from the pool of sexuals. With probability $q_p/(\eta+q_h+q_p)$, a mated slavemaker queen is picked from the pool of queens on their nuptial flights, and she searches the grid for a host nest to invade. A total number of n_p grid points are chosen, until a host nest is found. If none of these n_p grid points harbors a host nest, then no new slavemaker nest is founded. Otherwise, the first found host nest is invaded by the slavemaker queen and all host queens and adult workers in the nest are removed. The offspring is kept alive and will form the basis of the slave force. In either case, i.e., if a new slavemaker nest is founded or not, a slavemaker queen is removed from the pool of sexuals. If a host or slavemaker queen is chosen to attempt the foundation of a new nest, this does not count as a time step. The pool of sexuals is represented by double ended queues for host queens and males. Newly added sexuals are attached to the end of the queue. If a sexual is removed from a pool, it is taken from the front of the queue. In the simulation, there is no distinction between slavemaker queens, because their genotypes are not tracked. Hence, the pool of slavemaker queens is represented by an integer, which gets increased or decreased if queens are added or removed.

In a time step, a random grid point is chosen. If no nest is present, then nothing happens. If there is a host nest, then its (annual) actions are performed. First, if there are no workers present and the nest is at most one year old, then the current number of workers is set to five. Afterwards, if there are less than h_{\max} workers in the nest, and the nest has not recently been raided by slavemakers, then additional offspring is produced. Let h denote the current number of host workers in the nest. Then the number of new worker offspring is drawn from a Poisson distribution with mean $\lambda_h \cdot h \cdot (1 - h/h_{\max})$. If rebels are present in the nest, each newly produced worker is removed with probability μ_k . After production of offspring, if there are less than h_1 workers present, then the whole nest is removed with probability μ_{h1} . If there are more than h_2 workers, then the whole nest is removed with probability μ_{h2} . If there are less than h_2 and more than h_1 workers in the nest, then it gets removed with probability $\mu_{h1} + (h - h_1) \cdot \mu_{h2} - \mu_{h1} / h_2 - h_1$, where h denotes the number of workers currently in the nest. In case the nest is not removed, each worker dies with probability μ_h . If there is at most one worker left alive in the nest and the nest is not older than one year, or if there are at most three workers left alive in the nest and the nest is older than one year, the whole nest gets removed. Otherwise, if the nest has more than $h_{x \min}$ workers it produces a random number

of queens, according to a Poisson distribution with mean $\lambda_{xh} \cdot (h - h_{x \min})$. If there are rebels in the nest, each newly produced queen is killed with probability μ_k . The remaining queens and equally many males are added to the pool of sexuals currently on their nuptial flights. With a probability of p_{qr} , each of these queens returns to her home nest after mating with a male from the pool of sexuals. That male and the returning queen are removed from the pool of sexuals and the mated queen is included in the nest, leading to a polygynous nest. Then, each queen in the nest dies with probability $1/(1+e^{-(a-a_h)b_h})$, where a is the age of the queen. If this renders the nest queenless, it is henceforth treated as a satellite nest, as described below, and the nest performs no further actions in this simulation step. Otherwise, the age of each surviving queen is increased by one. The age of the nest corresponds to the age of the oldest queen. If there are more than $h_{s \min}$ workers present in the nest, a satellite nest is built with probability $\min\{1, (h - h_{s \min}) \cdot \zeta\}$. If that happens, two random numbers are drawn independently from a normal distribution with mean zero and standard deviation d_b . They are rounded to integers and added to the x and y coordinates of the current nest, respectively. If at the resulting coordinates (taken modulo 200, to account for the torus structure) there is no nest, the satellite nest is built there. Otherwise, the procedure is repeated up to a total of ten times. If no vacant grid point can be found, then the satellite nest does not get built. In the case that a vacant spot is found, a new nest will be established there. With probability p_{sh} , or if the pool of males is empty, this nest does not get a queen. Otherwise it gets a queen, who stems from the mother nest, i.e., she inherits the genotype from a randomly drawn (mated) queen of the mother nest. The new queen mates with a male from the pool of sexuals. This male then gets removed from the pool. Independent of the new nest having a queen or not, each worker from the mother nest joins the satellite nest with probability u_h . The nests then belong to the same colony and exchange workers. This does not affect the simulation run, but when relatedness values are calculated for the summary statistics, and ants are sampled from the nests, each sampled ant is chosen from another nest of the same colony (if the colony is larger than one nest) with probability p_{hx} .

Queenless host satellite nests are treated like nests with a queen, except for the following differences. They are removed if they become older than three years, i.e., if they are selected to perform their annual actions for the fourth time. Furthermore, they do not found satellite nests to extend the colony and they do not produce females, i.e., workers or queens, but only males. The number of males produced is Poisson distributed with mean $2 \cdot \lambda_{xh} \cdot (h - h_{x \min})$, where h is current worker count of the satellite nest. These males are then added to the pool of sexuals.

When a slavemaker nest is selected in a time step to perform its actions and if the number of slaves and slavemaker workers combined is less than w_{\max} , then new workers are pro-

duced. The amount of new workers is Poisson distributed with mean $s \cdot \lambda_s \cdot (1 - s+p/w_{\max})$, where s and p denote the numbers of slaves and slavemaker workers in the nest, respectively. If there are rebels among the slaves in the nest, then each newly produced slavemaker worker is killed with probability μ_r . Then, with probability $\max\{1, \min\{0, (1 - s/s_{\max}) \cdot \mu_{p1}\}\}$, the whole nest gets removed. In case the nest does not get removed, each slavemaker worker dies with probability μ_p , if there are at most p_{\max} slavemaker workers in the nest. Otherwise, this dying probability is increased by $p-p_{\max}/p$, with p denoting the number of slavemaker workers. Similarly, if there are at most s_{\max} slaves in the nest, each slave dies with probability μ_s . If more slaves are present, this is increased by $s-s_{\max}/s$, where s is defined as the number of slaves in the nest. Genotypes of the slaves are stored in the slavemaker nest. To that end, for each host nest that slaves in the slavemaker nest originate from, the genotypes of all queens and males are stored, together with the number of slaves that this host nest contributes to the current slave force. The choice of which slaves get removed in case of death is performed uniformly among all slaves.

If there are more than $s_{x \min}$ slaves left in the nest, then a random number of new slavemaker queens is produced, according to a Poisson distribution with mean $\lambda_{xs} \cdot (s - s_{x \min})$, where s denotes the number of slaves in the nest. If there are rebels among the slaves, then each of the newly produced queens is killed with probability μ_r . The surviving ones are then added to the pool of sexuals, currently on their nuptial flight. Slavemaker males are not considered in the simulation, because slavemaker genotypes do not play a role.

If there is at least one slavemaker worker in the nest, then a Poisson distributed number of raids, with mean $p \cdot n_r$, is performed, where p denotes the number of slavemaker workers in the nest. To determine which nest is raided, two random numbers are drawn independently from a normal distribution with mean zero and standard deviation d_r . They are rounded to integers and added to the x and y coordinates of the current nest, respectively. If at the resulting coordinates (taken modulo 200, to account for the torus structure) there is a host nest, then it gets raided. Otherwise, the procedure is repeated up to a total of ten times. If no host nest is found, then no raid is performed. If a host nest with a queen is found, then the number of host worker offspring is determined, as described above. If the nest survives the raid, then it will not produce further offspring the next time it is selected to perform its actions. Each newly produced host worker offspring is taken by the raiding slavemakers with probability f_r . Furthermore, during the raid, each host queen, each adult host worker, and each adult slavemaker worker is killed with probabilities ξ_f , $1 - \alpha_h$, and ρ , respectively. If there are rebels in the slavemaker nest, then each of the newly captured slaves is killed before reaching adulthood with probability μ_q . In case that there are no host queens remaining in the raided nest, either because there were none present prior to the raid, or because they were

all killed during the raid, and if there are more than five slavemaker workers in the current slavemaker nest, the host nest has a chance of being captured by the raiding slavemakers and turned into a satellite nest. The probability of this happening is $\max\{1/2 + p^{-6}/20, 1\}$, where p denotes the number of slavemaker workers in the current slavemaker nest. If a satellite nest is built, then $\max\{4, \lfloor p/3 \rfloor\}$ (where $\lfloor \cdot \rfloor$ denotes the “floor function”) of the slavemaker workers from the originating slavemaker nest stay there and the satellite nest gets a new slavemaker queen with probability p_{pq} . If the raided nest is not turned into a slavemaker nest and no surviving host workers are left, the nest is removed. After all raids are performed, if there are less than four slaves in the slavemaker nest and it is older than a year, or if there is at most one slave and the slavemaker nest is not older than one year, then the slavemaker nest is removed. With probability $1/(1+e^{-(a-ap)\cdot bp})$, the slavemaker queen dies, where a denotes her age. If the queen survives, her age is increased by one. Otherwise, the slavemaker nest is rendered queenless. If a queenless slavemaker nest is picked to perform its actions and it is older than three years, then it is removed. It cannot produce workers or queens and it is not removed if it has less than four slaves after performing raids. Apart from these differences, queenless slavemaker nests are treated in the same way as slavemaker nests with a queen. Unlike host nests, slavemaker nests are never polygynous in the simulation model.

After a time step has been performed, a random number of host queens and males and slavemaker queens immigrate to the area and are added to the pool of sexuals. These numbers are chosen according to a Poisson distribution with mean $m_h/40\,000$ for hosts, independently for queens and males. For slavemaker queens, this number is Poisson distributed with mean $m_p/40\,000$. Immigrating host queens and males are assigned random genotypes with probability f_A of having the rebel allele on a given chromosome. Microsatellites are distributed according to empirical field observations, matching the frequencies of the initial simulation setup. After the immigration, random numbers of host queens and males are removed from the front of the double ended queues, representing the pools of sexuals. This number is binomially distributed with parameters q_h and μ_{xh} , where q_h denotes the number of host queens in the pool, for queens, and with parameters m and μ_{xh} , where m denotes the number of host males in the pool, for males. Furthermore, a random number of slavemaker queens is removed from the pool, according to a binomial distribution with parameters q_p and μ_{xp} , where q_p denotes the current number of slavemaker queens in the pool.

<i>Param.</i>	<i>Prior</i>	<i>Meaning</i>
f_A	$\{0, \mathcal{U}(0, 1)\}$	Rebel allele frequency in the beginning and in migrants
h_{\max}	$\ln \mathcal{N}(80, 15)$	Capacity for host workers in host nests
s_{\max}	$\ln \mathcal{N}(80, 15)$	Capacity for slaves in slavemaker nests

p_{\max}	$\ln \mathcal{N}(20, 5)$	Capacity for slavemaker workers in slavemaker nests
w_{\max}	$\ln \mathcal{N}(80, 15)$	Capacity for sum of slavemaker workers and slaves in slavemaker nest
λ_h	$\ln \mathcal{N}(2, 1.5)$	Relative productivity of host workers
λ_s	$\ln \mathcal{N}(2, 1)$	Relative productivity of slaves
μ_r	Beta(0.8, 0.125)	Fraction of slavemaker offspring that is killed when rebels are present
μ_k	Beta(0.05, 0.05)	Fraction of host offspring in free nest that is killed when rebels are present
μ_q	Beta(0.3, 0.15)	Fraction of captured slave offspring that is killed by rebels in a slavemaker nest
μ_h	Beta(0.3, 0.1)	Death prob. of a host worker in host nest per year
μ_s	Beta(0.3, 0.1)	Death prob. of a slave per year
μ_p	Beta(0.3, 0.1)	Death prob. of a slavemaker worker per year
$h_{x \min}$	$\ln \mathcal{N}(13, 3)$	Minimum number of workers needed to produce queens and males in hosts
$s_{x \min}$	$\ln \mathcal{N}(13, 3)$	Minimum number of slaves needed to produce queens and males in slavemakers
λ_{xh}	$\ln \mathcal{N}(0.25, 0.1)$	Increase of number of produced sexuals per worker in hosts
λ_{xs}	$\ln \mathcal{N}(0.25, 0.1)$	Increase of number of produced sexuals per slave in slavemakers
p_{qr}	Beta(0.005, 0.005)	Prob. of a host queen to return to her nest after mating
h_1	$\ln \mathcal{N}(5, 2)$	Upper bound for number of workers in host nest such that the annual death prob. of the nest is μ_{h1}
μ_{h1}	$\ln \mathcal{N}(0.1, 0.05)$	Prob. of host nest death if it has at most h_1 workers
h_2	$\ln \mathcal{N}(40, 10)$	Lower bound for number of workers in host nest such that the annual death prob. of the nest is μ_{h2}
μ_{h2}	Beta(0.02, 0.01)	Prob. of host nest death if it has at least h_2 workers
a_h	$\ln \mathcal{N}(25, 5)$	Age of host queen with annual dying prob. 0.5
a_p	$\ln \mathcal{N}(25, 5)$	Age of slavemaker queen with annual dying prob. 0.5
b_h	Beta(0.45, 0.12)	Influence of age to dying prob. of host queen
b_p	Beta(0.45, 0.12)	Influence of age to dying prob. of slavemaker queen

μ_{p1}	Beta(0.1, 0.08)	Maximum annual dying prob. for slavemaker nests
$h_{s\min}$	$\ln \mathcal{N}(35, 5)$	Minimum number of workers required for a host nest to be able to build a satellite nest
ζ	Beta(0.04, 0.02)	A host nest with $h > h_{s\min}$ workers has prob. $\min\{1, (h - h_{s\min})\zeta\}$ to build a satellite nest
d_b	$\ln \mathcal{N}(2, 1.1)$	Branching distance of host nests
p_{sh}	Beta(0.8, 0.1)	Prob. that a newly founded satellite host nest does not get a queen
u_h	Beta(0.3, 0.1)	Fraction of host workers that move to a newly founded satellite nest
p_{hx}	Beta(0.1, 0.1)	Proportion of workers that is exchanged between nests of a host colony
n_r	$\ln \mathcal{N}(0.5, 0.15)$	Expected annual number of raids per slavemaker worker
d_r	$\ln \mathcal{N}(2.7, 1.2)$	Raiding distance
f_r	Beta(0.65, 0.15)	Fraction of host offspring that is raided
α_h	Beta(0.75, 0.15)	Fraction of host workers in the raided nest that survive
ρ	Beta(0.05, 0.05)	Fraction of slavemaker workers that are killed on a raid
ξ_f	Beta(0.9, 0.1)	Prob. that a host queen dies while her nest is raided
p_{pq}	Beta(0.1, 0.05)	Prob. that a satellite slavemaker nest founded during raid gets a queen
μ_{xh}	$\mathcal{U}(0, 1)$	Prob. for each host individual in the pool to die or emigrate in one time step
μ_{xp}	$\mathcal{U}(0, 1)$	Prob. for each slavemaker queen in the pool to die or emigrate in one time step
η	$\ln \mathcal{N}(250, 200)$	Determines the rate at which new nests are founded
n_h	$\ln \mathcal{N}(10, 5)$	Number of sites a host queen checks to find a free one, when attempting to found nest
n_p	$\ln \mathcal{N}(10, 5)$	Number of sites a slavemaker queen checks to find a host nest, when attempting to invade a nest
m_h	$\ln \mathcal{N}(20, 20)$	Mean number of immigrating host queens and males per year
m_p	$\ln \mathcal{N}(2, 2)$	Mean number of immigrating slavemaker queens per year

Table 2.4: Prior distributions for ABC analysis. Table adapted from Metzler et al. (2015).

Box 2.7: Simulation scheme. Parameters are described in Table 2.4 and Table 2.6.

```
Set  $t := 0$ 
while  $t < T$  do
  Draw uniform real number  $rand \in (0, \eta + q_h + q_p)$ 
  if  $rand < \eta$  then
    Choose random grid point
    if there is a host or slavemaker nest then
      Perform nest actions (see Box 2.8 and Box 2.9)
    end if
    Host and slavemaker sexuals immigrate to the pool
    Random deaths of sexuals in the pool
    Set  $t := t + \frac{1}{N^2}$ 
  else if  $rand < \eta + q_h$  then
    Set  $trial := 0$ 
    do
      Choose random grid point
      Set  $trial := trial + 1$ 
      while grid point not empty &  $trial < n_h$ 
      if empty site found then
        Host queen from pool builds nest at empty site
      else
        Remove a host queen and male from the pool
      end if
    else
      Set  $trial := 0$ 
      do
        Choose random grid point
        Set  $trial := trial + 1$ 
        while no host nest at grid point &  $trial < n_p$ 
        if host nest found then
          Slavemaker queen from pool invades host nest
        else
          Remove a slavemaker queen from the pool
        end if
      end if
    end if
  end while
```


Box 2.8: Simulation scheme for annual host nest actions. See Table 2.4, Table 2.5, and Table 2.6 for a description of the parameters and variables.

```

if  $h < h_{\max}$  then
  if  $h = 0$  & age of nest  $< 2$  then
    Set  $h := 5$ 
  end if
  Produce offspring
end if
Whole nest may be removed with probability depending on
current nest size
Each worker dies with probability  $\mu_h$ 
if  $h \leq 1$  | ( $h \leq 3$  & age of nest  $> 1$ ) then
  Remove nest
end if
Produce queens and males and send them in the pool of
sexuals
Each queen dies with probability depending on her age
if  $h > h_{s\min}$  then
  A satellite nest may be built with probability
  depending on current nest size
end if

```

Box 2.9: Simulation scheme for annual slavemaker nest actions. See Table 2.4, Table 2.5, and Table 2.6 for a description of the parameters and variables.

```

if  $p + s < w_{\max}$  then
  produce offspring
end if
Whole nest may be removed with probability depending on
current nest size
Each slavemaker worker dies with probability  $\mu_p$ 
(increased if  $p > p_{\max}$ )
Each slave dies with probability  $\mu_s$  (increased if  $s > s_{\max}$ )
Produce queens and send them in the pool of sexuals
A Poisson( $p \cdot n_r$ )-distributed number of raids on host nests
in the vicinity is performed
if  $s \leq 1$  | ( $s \leq 3$  & age of nest  $> 1$ ) then
  Remove nest
end if
Whole nest may be removed with probability depending on
current nest age

```

<i>Variable name</i>	<i>Meaning</i>
h	Number of workers in current host nest
s	Number of slaves in current slavemaker nest
p	Number of slavemaker workers in current slavemaker nest
a	Age of a queen in the current nest
q_h	Number of host queens currently on their nuptial flights.
q_p	Number of slavemaker queens currently on their nuptial flights.

Table 2.5: Names of variables as used in the specification of the simulation model.

Parameter	Description	Value
T	Time horizon in years	5000
N	Number of grid points on horizontal and vertical axis	200
H_0	Expected value of number of host nests on the grid at the beginning of the simulation	8750
P_0	Expected value of number of slavemaker nests on the grid at the beginning of the simulation	875

Table 2.6: Model parameters that are fixed across different simulation studies, together with the respective values. See Table 2.4 for parameters chosen according to prior distributions or samples from posterior distributions.

2.3.2 Simulation study

In this section, the various simulation series of the simulation study in Metzler et al. (2015), investigating the evolution of slave rebellion in ants, are described. A simplified overview is given in Figure 2.1.

2.3.2.1 Prior simulations

The simulation model incorporates a lot of biologically relevant aspects of the system of ant slavemakers and their hosts. Although the ecology of *Protomognathus americanus* and *Temnothorax longispinosus* is studied intensively (Herbers, 1989; Foitzik and Herbers, 2001a,b; Herbers and Foitzik, 2002; Brandt and Foitzik, 2004; Foitzik et al., 2004; Fischer and Foitzik, 2004; Brandt et al., 2005a,b; Fischer-Blass et al., 2006; Brandt et al., 2007; Foitzik et al., 2009; Achenbach and Foitzik, 2009; Achenbach et al., 2010; Pohl and Foitzik, 2011; Pennings et al., 2011; Scharf et al., 2011; Konrad et al., 2012; Pamminger et al., 2011, 2012, 2013, 2014), it is still difficult to exactly determine parameter values that resemble the situation in the field. Furthermore, ecological parameters can vary across geographic locations. To account for the complexity of the system, the simulation model has 47 parameters. This leads to a huge parameter space and a crucial first step is to find appropriate parameter

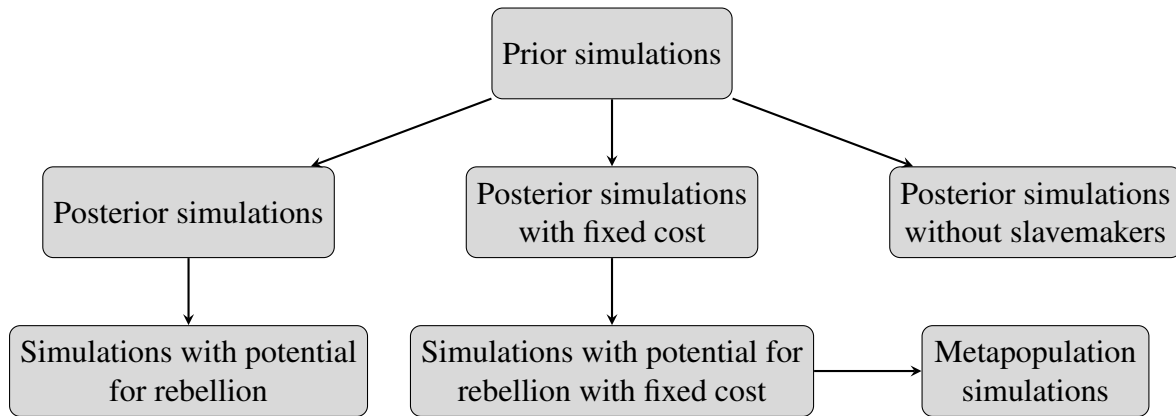


Figure 2.1: Overview of the simulation study to investigate the evolution of slave rebellion in ants. Arrows indicate which simulation series was used to determine parameters of subsequent simulation series.

ranges. Many of the parameters relevant for our model are extremely difficult to obtain from field observations, as they would require constant monitoring of all ants in the area. Nevertheless, the existing studies help narrowing down the parameter space. The resulting prior distributions are provided in Table 2.4 for all parameters. Further reduction of the parameter ranges can be obtained using the Approximate Bayesian Computation (ABC) methodology (see Csilléry et al., 2010, for a review). In the first simulation series (*Prior simulations*), all parameters were drawn independently from the specified prior distributions to perform 10 000 simulation runs. For half of these runs, the rebel allele was absent. In the other half, the rebel allele frequency was drawn uniformly from the interval $(0, 1)$. In the ABC approach, the generated output was compared to field data from New York State and West Virginia, gathered by Pamminer et al. (2014). This comparison was achieved using 14 summary statistics. These are average number of host workers in host nests, average number of slaves in slavemaker nests, fractions of queenless and polygynous nests among host nests, relatedness of host workers within host nests, relatedness of host workers between host nests at distances of at most 0.8 m, at distances between 0.8 m and 1.75 m, and at distances between 1.75 m and 3 m, relatedness of slaves within slavemaker nests, relatedness of slaves and free-living host workers between slavemaker nests and host nests at distances of at most 0.8 m, at distances between 0.8 m and 1.75 m, and at distances between 1.75 m and 3 m. Relatedness values were calculated from six microsatellite loci using equations (6) and (10) of Queller and Goodnight (1989). All of the above summary statistics refer to nests with at least five host workers for host nests and at least five slaves for slavemaker nests. The remaining two summary statistics are number of host and slavemaker nests per square meter. A list of all summary statistics with observed values from the field data gathered in New York and

Summary statistics	Observed NY	Observed WV
Avg. number of host workers	21.04	18.03
Avg. number of slaves	26.9	27.32
Fraction of queenless host nests	0.3756	0.3319
Fraction of polygynous host nests	0.0335	0.0596
Number of host nests per m ²	0.9	0.6
Number of slavemaker nests per m ²	0.15	0.1
Relatedness within nests	0.4614	0.5879
Relatedness between nests, $\text{dist} \leq 0.8$ m	0.1217	0.2704
Relatedness between nests, $0.8 \text{ m} < \text{dist} \leq 1.75$ m	0.0536	0.0738
Relatedness between nests, $1.75 \text{ m} < \text{dist} \leq 3$ m	0.0111	0.0217
Relatedness of slaves within nests	0.2578	0.3612
Rel. slaves/free ants betw. nests, $\text{dist} \leq 0.8$ m	0.1224	0.1827
Rel. slaves/free ants betw. nests, $0.8 \text{ m} < \text{dist} \leq 1.75$ m	0.0481	0.0145
Rel. slaves/free ants betw. nests, $1.75 \text{ m} < \text{dist} \leq 3$ m	0.0293	0.0635

Table 2.7: Summary statistics used in the ABC analysis together with observed values from New York (NY) and West Virginia (WV) by Pamminger et al. (2014). Table is adapted from Metzler et al. (2015).

West Virginia by Pamminger et al. (2014) is provided in Table 2.7. The summary statistics were chosen to be independent of the existence of rebellion behavior. The ABC analysis was performed with the *R*-package *abc* (R Core Team, 2015; Csilléry et al., 2012) and the local-linear regression method by Beaumont et al. (2002) was used. This was done separately for the New York and the West Virginia data. For each of the two data sets, among all of the 10 000 simulation runs, those 5% were kept that produced summary statistics most similar to the observed ones under Euclidean distance. The corresponding 500 parameter sets per location were retained as samples of the posterior distributions.

2.3.2.2 Posterior simulations

For each of the 500 parameter sets per location, obtained as samples from the posterior distributions in the *Prior simulations*, ten independent simulations were performed. Hence, this simulation series (*Posterior simulations*) consisted of 10 000 simulation runs. The cost parameter describing the influence of the rebel allele in free-living host nests ranged from values as low as 0.0003 to values as large as 0.6818. The runs in the *Prior simulations* accounted for various levels of frequency of the rebellion allele as well as its complete absence. Furthermore, the ABC analysis was performed without conditioning on the existence of the rebellion allele. The goal of this simulation series was to find conditions under which the

rebel allele can spread in the population, starting at a low frequency. Thus, the frequency of the rebel allele in the beginning of the simulation and in immigrating host ants was fixed to 0.01. First, all simulations that led to a final host nest density of less than 0.1 per m² were excluded. Then, to check the success of the rebellion allele, its frequency at the end of each simulation run was computed.

Furthermore, the parameter combinations were checked for their potential of showing an increased rebel allele frequency over a long period of time. In all simulations, allele counts were stored only at the end of each simulation run, to keep the runtime of the program low. Thus, for the long-term calculations of rebel frequencies, nest counts were used as a proxy. A host nest was counted as a rebel nest, if at least one parent of the nest had at least one rebel allele. Otherwise, it was counted as a non-rebel nest. Then, the average numbers of rebel host nests and non-rebel host nests per year in the last 1000 years of each simulation were computed. From these, the ratio of the average number of rebel nests divided by the average number of all host nests (i.e., the sum of the averages for rebel and non-rebel nests) was calculated.

2.3.2.3 Posterior simulations with fixed cost

The cost parameter determines the probability that a host pupae in a free-living host nest containing bearers of the rebel allele is erroneously killed before reaching adulthood. A larger value of the cost parameter means that rebel host nests produce fewer adult queens, males, and workers, and therefore have a stronger disadvantage compared to non-rebel host nests. Thus, the cost parameter is likely to have a strong impact on the possibility of the rebel allele reaching a high frequency and successfully establishing in the host population. In the runs of the *Posterior simulations*, the cost parameter varied between 0.0003 and 0.6818. Extremely large values can completely rule out the possibility of the rebel allele evolving, while very low values may mean that the rebel allele bears practically no cost. In this simulation series, the goal was to find conditions for the rebel allele to evolve, when it is costly, without the cost being extremely large. Thus, in this simulation series (*Posterior simulations with fixed cost*), two scenarios of a costly rebel allele were explored and the parameter was fixed to 0.01 and 0.05. Apart from that, the same parameter sets as for the *Posterior simulations* were used. For both cost settings, for each of the 500 parameter sets of both locations, three simulation runs were performed. This yields a total of 6000 simulation runs. For these, long-term rebel nest frequencies over the last 1000 simulated years were computed, similar as for the *Posterior simulations*.

2.3.2.4 Posterior simulations without slavemakers

The benefit that the rebel allele brings to the host population lies in the reduction of slave-maker pressure. Thus, the rebel allele should not be able to reach high frequencies, if there are no slavemakers in the area. In order to test this hypothesis, slavemakers were excluded entirely in this simulation series (*Posterior simulations without slavemakers*). Otherwise, the same parameter sets as those for the *Posterior simulations* were used. For both locations—New York and West Virginia—for each of the 500 parameter sets, three simulations were performed, leading to a total of 3000 simulation runs.

To compare the results to the simulations with slavemakers from Section 2.3.2.2, those runs with a final host nest density below 0.1 per m² were excluded. Then, for each parameter set, the mean value of the final rebel allele frequency was computed from the remaining runs. In the same way, the mean final rebel allele frequency was computed for the first three of the ten simulations per parameter set from the *Posterior simulations*. Parameter sets leading to a mean final rebel allele frequency below 10^{-5} in both cases—with and without slavemakers—were excluded from the further analysis. With the remaining parameter sets, the null hypothesis that it is equally likely that simulations with or without slavemakers lead to a higher final rebel allele frequency, was tested.

2.3.2.5 Simulations with potential for rebellion

In this simulation series (*Simulations with potential for rebellion*), those parameter sets from the *Posterior simulations* were examined that showed the potential for an increased rebel frequency over a long time period. To that end, parameter combinations from the *Posterior simulations* that led to at least one simulation for which the rebel nest frequency averaged over the last 1000 years of the simulation was above 5%, were selected. For each of the selected parameter sets, 20 additional simulation runs were performed, with the goal of checking whether the increased rebel frequency over a long time-period was reproducible.

2.3.2.6 Simulations with potential for rebellion with fixed cost

The choice of parameter sets for this simulation series (*Simulations with potential for rebellion with fixed cost*) was performed in an analogous manner to the choice for the *Simulations with potential for rebellion*, but based on the simulations of the *Posterior simulations with fixed cost*. For all parameter combinations of the *Posterior simulations with fixed cost* that led to an increased average rebel nest frequency over the last 1000 years of the simulation, as described in Section 2.3.2.2 for the *Simulations with potential for rebellion*, 40 additional simulations were performed. To check whether the increased rebel frequency over a long

time-period found in the parameter sets of the *Posterior simulations with fixed cost* was reproducible, the current simulations were checked for the two criteria that (a) during the last 1000 years of the simulation, the average host nest density was larger than 0.1 per m² and (b) during the last 1000 years of the simulation, the rebel allele frequency among all produced host queens and males was larger than 0.01.

Furthermore, to test for deme-level effects, the average rebel allele frequency among all produced host queens and males during the last 1000 years over all 40 simulation runs was computed for all 37 parameter sets.

To investigate effects of single parameters, all parameter combinations that showed an increased rebel allele frequency among produced sexuals, as well as a minimum nest density of 0.1 per m², were compared to a random sample of all 1000 parameter combinations of the *Posterior simulations with fixed cost*. All parameters were compared individually using a Wilcoxon rank test. The Bonferroni-Holm correction was applied to account for multiple testing.

2.3.2.7 Metapopulation simulations

In a given simulation, a high rebel nest frequency may protect the host population from immigrating slavemakers. This could result in an increased productivity of host nests, which in turn might lead to a high number of emigrating host queens and males. These would then be likely to carry the rebel allele as well and thus, spread it to other locations. In order to assess the possibility of this scenario, a metapopulation of 20 demes was simulated in this simulation series (*Metapopulation simulations*). Each deme was represented by the grid of 200×200 points, as described in Section 2.3.1. Here, a time span of 5200 years was simulated. During the first 200 years, the frequency of the rebel allele among all immigrating host queens and males was set to 0.01. After this burn-in time, for the remaining 5000 years, the frequency of the rebel allele among immigrating host queens and males was set to the average rebel allele frequency among all host queens and males produced in the 20 demes. In case that the host population went extinct in all 20 demes, the rebel allele frequency among immigrants was again set to 0.01. The rate of immigration was set to the value of the corresponding parameter combination, and thus was independent of the productivity of the metapopulation. Ten independent simulations were performed per parameter set. Among the parameter combinations of the *Simulations with potential for rebellion with fixed cost*, those were chosen for the present simulation series, which fulfilled both criteria—(a) and (b), described in Section 2.3.2.6—in more than a quarter of the 40 simulation runs.

In each simulation, every ten years, for all host nests with coordinates (i, j) , such that

$i, j \in \{3, 8, 13, 18, \dots, 198\}$ (i.e., all nests at the center of disjoint blocks of 5×5 grid points) fitness and rebel allele frequencies among queens and fathers of the nests were recorded. Nest fitness here was defined as the sum of the number of males and twice the number of queens, produced by the nest within the following ten years (or until the nest died, if that occurred earlier). From these values, the within-deme and between-deme components of Queller's group selection variant of the Price equation (equation (22) in Queller, 1992; see also Price, 1970) could be approximated. The within-deme component was calculated as the average of the covariance of relative nest fitness (i.e., nest fitness divided by the sum of all nest fitness values in the deme) and allele frequency, taken over all 20 demes, weighted by the number of host nests found at the coordinates specified above in a deme (which acts as a proxy for the deme's host population density). The between-deme component was computed as the covariance of the rebel allele frequency in the whole deme and the deme-wide average relative nest-fitness.

Chapter 3

Results

3.1 Mathematical analysis and asymptotics

The results presented in this section are based on Hutzenthaler et al. (2015). A condition for fixation or extinction of an altruistic defense allele is obtained, starting from an individual-based model of populations structured in demes. The diffusion approximation of this model is derived and after taking the limits of infinitely large populations, infinitely many demes, and an infinite time horizon, the result is reached. This result states that the allele for altruistic defense will fixate or go extinct, depending on the ecological model parameters. A stable equilibrium between altruists and cheaters is reached only for the critical values of the model parameters. An overview of the steps involved in the derivation is given in Figure 3.1. The technical details as well as the proofs can be found in Chapter 4. A simple individual-based model for the evolution of an altruistic defense allele is specified in Section 2.1.1. The numbers of individuals of the different types—altruists, cheaters, and parasites—structured in demes form the states of a Markov chain. Transition rates are given by (2.1). This simple model incorporates random birth and death events, interactions between hosts and parasites, intra-species competition, migration, and the effect of the altruists. Namely, altruists provide help to all hosts by lowering the growth rates of parasites in their deme, while at the same time suffering the cost of having an increased death rate compared to cheaters. In order to analytically handle this model, its diffusion approximation is derived in Section 4.1.2. The resulting diffusion model, given by equation (4.9) in Section 4.1.3, is an extension of a classical Lotka-Volterra equation (Lotka, 1920; Volterra, 1926) with stochasticity, spatial structure, and altruistic defenders. This model is analyzed in Sections 4.1.4 to 4.1.7. There are added immigration terms in the diffusion equations— $\int_0^t \iota_H^N \frac{A_s^N(i)}{A_s^N(i) + C_s^N(i)} ds$, $\int_0^t \iota_H^N \frac{C_s^N(i)}{A_s^N(i) + C_s^N(i)} ds$, and $\int_0^t \iota_P^N ds$ —which, if they are large enough, prevent the host and

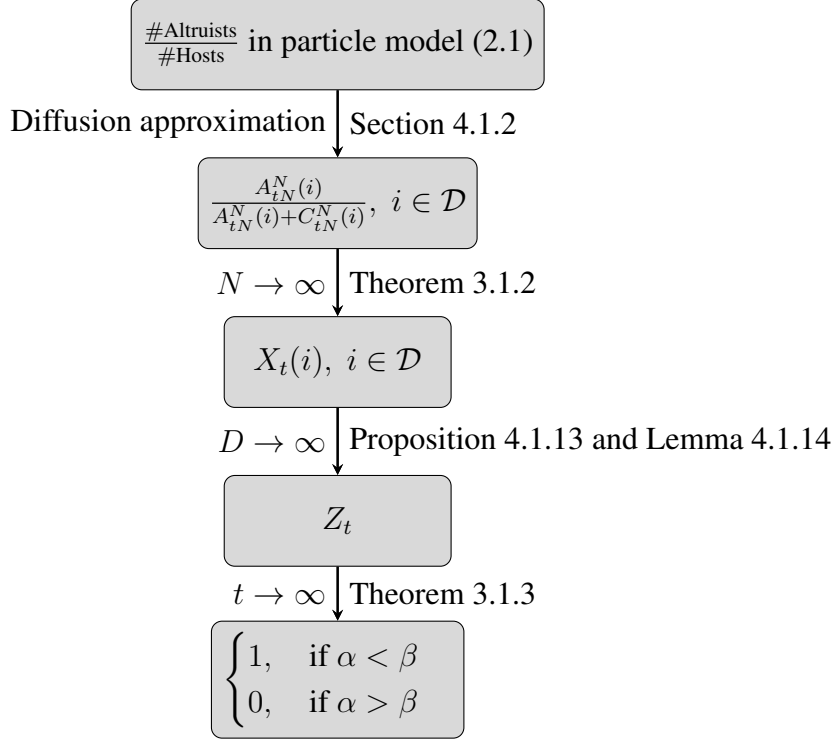


Figure 3.1: Simplified overview of the results from the theoretical analysis on the relative frequency of an altruistic defense allele in a structured host population.

parasite populations from local extinction due to random fluctuations, in the finite setting $N < \infty$, as shown in Lemmas 4.1.3 and 4.1.4. Note that either altruists or cheaters may die out, while the total host population is prevented from going extinct due to the immigration terms. First, the equations for altruists, cheaters, and parasites are transformed, to obtain the equivalent equations (4.14), (4.15), and (4.16), describing the total number of host individuals, the relative altruist frequency among all hosts, and the number of parasite individuals, respectively. The goal is to understand the evolution of the altruistic defense allele. Thus, the dynamics of the process F^N that solves equation (4.15) is analyzed. For all $N \in \mathbb{N}$, the process F^N still depends on the process H^N , which in turn depends on P^N , and the first results in Section 4.1.4 are concerned with the behavior of the latter two processes. Theorem 4.1.9 extends the results on a well-known Lyapunov function (e.g., Dobrinevski and Frey, 2012) to the spatial setting. Its proof rests on Lemmas 4.1.6, 4.1.7, and 4.1.8, which show boundedness of the expected value of weighted sums of functionals of H^N and P^N . Intuitively, these lemmas show that under certain assumptions, the interactions of H^N and P^N keep (the expected value of) either process, as well as certain combinations of the two, from becoming “too large” or “too small”. As these processes are strongly connected, it does not suffice to investigate them separately. Instead, it is necessary to check their behavior by combining

them in a suitable way, as is done in the aforementioned Lemmas. Proofs of Lemmas 4.1.6, 4.1.7, and 4.1.8 and Theorem 4.1.9 are given in Section 4.1.4.3. An immediate consequence of Theorem 4.1.9, after using a time substitution, is Theorem 3.1.1. This result formalizes the effect of a complete separation of time scales in the present scenario. The assumptions of weak selection, weak migration, and weak genetic drift ensure that evolutionary forces act much slower than ecological forces. Under this separation of time scales, the relative altruist frequency F^N , which in a large population is influenced only by evolutionary forces, stays constant over long periods of time, while numbers of hosts and parasites change more rapidly, due to ecological forces. Thus—as in the case of classical Lotka-Volterra equations (Lotka, 1920; Volterra, 1926)—an equilibrium between hosts and parasites is reached after a period of fluctuations, but before any changes in F^N can be seen. By speeding up the processes appropriately, numbers of hosts and parasites will immediately be at their equilibrium values and stay at these values—which depend on F^N —at all times. The mathematically rigorous statement of this insight is given in Theorem 3.1.1.

Theorem 3.1.1. *Assume the setting of Section 4.1.3, let Assumption 4.1.1 hold, assume that $\sup_{N \in \mathbb{N}} (N \max\{\kappa_H^N, \kappa_P^N, \alpha^N, \iota_H^N, \iota_P^N, \beta_H^N, \beta_P^N\}) < \infty$ and let h_∞ and p_∞ be given by (4.13). Then we get for all sets $\hat{\mathcal{D}} \subseteq \mathcal{D}$ and all $t \in [0, \infty)$ that*

$$\sup_{N \in \mathbb{N}} N \int_0^t \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(H_{uN}^N(i) - h_\infty(F_{uN}^N(i)) \right)^2 + \sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(P_{uN}^N(i) - p_\infty(F_{uN}^N(i)) \right)^2 \right] du < \infty. \quad (3.1)$$

To the best of my knowledge, this result is new even in a panmictic setting, and apart from proving for every $t \in [0, \infty)$ the $L^2([0, t] \times l_\sigma^1 \times \Omega; \mathbb{R})$ -convergence of $(H_{\cdot N}^N, P_{\cdot N}^N)$ to $(h_\infty(F_{\cdot N}^N), p_\infty(F_{\cdot N}^N))$ as $N \rightarrow \infty$, it also shows the rate to be at least $1/2$. Theorem 3.1.1 states that, when all processes are sped up by a factor of N , then H^N and P^N can be approximated for large N by functions h_∞ and p_∞ of F^N , as given by equation (4.13). Using this approximation in equation (4.15) leads to the limit process of $(F_{\cdot N}^N)_{N \in \mathbb{N}}$, as $N \rightarrow \infty$. Theorem 3.1.2 shows this limit process X , which then describes the altruist frequency on the sped-up evolutionary time scale. Convergence in distribution is denoted by “ \implies ” (e.g., Klenke, 2008). Theorem 3.1.2 provides a simplified expression for the relative altruist frequency. This is achieved by making use of the separation of time scales together with the insight from Theorem 3.1.1 that numbers of hosts and parasites are always at their equilibrium values, when the processes are studied on the evolutionary time scale. Replacing these expressions then yields an equation for the altruist frequency, which no longer depends on

the numbers of hosts and parasites per deme and is thus easier to analyze.

Theorem 3.1.2. *Assume the setting of Section 4.1.3, let Assumption 4.1.1 hold, assume that $\sum_{i \in \mathcal{D}} \sup_{N \in \mathbb{N}} \sigma_i \mathbb{E}[H_0^N(i)] < \infty$, that $\sup_{N \in \mathbb{N}} (N \max\{\kappa_P^N, \iota_H^N, \iota_P^N, \beta_P^N\}) < \infty$, that there exist $\kappa, \alpha, \beta \in [0, \infty)$ such that $\lim_{N \rightarrow \infty} \kappa_H^N N = \kappa$, such that $\lim_{N \rightarrow \infty} \alpha^N N = \alpha$, and such that $\lim_{N \rightarrow \infty} \beta_H^N N b = \beta$ and assume that $F_0^N \Rightarrow X_0$ as $N \rightarrow \infty$ in l_σ^1 . Then the SDE*

$$\begin{aligned} dX_t(i) = & \kappa \sum_{j \in \mathcal{D}} m(i, j) \frac{a - X_t(i)}{a - X_t(j)} (X_t(j) - X_t(i)) dt - \alpha X_t(i)(1 - X_t(i)) dt \\ & + \sqrt{\beta(a - X_t(i))X_t(i)(1 - X_t(i))} dW_t(i), \quad t \in (0, \infty), i \in \mathcal{D} \end{aligned} \quad (3.2)$$

(where $\{W(i) : i \in \mathcal{D}\}$ are independent standard Brownian motions) has a unique strong solution and

$$(F_{tN}^N)_{t \in [0, \infty)} \Rightarrow (X_t)_{t \in [0, \infty)} \quad (3.3)$$

as $N \rightarrow \infty$ in $C([0, \infty), l_\sigma^1)$.

The proof, which is given in Section 4.1.4.4, builds on Theorem 3.1.1 and uses a stochastic averaging result from Kurtz (1992). Due to Theorem 3.1.2, in order to analyze the evolution of the altruistic defense allele, now only X has to be studied, instead of keeping track of A^N , C^N , and P^N (or, equivalently, of H^N , F^N , and P^N). The process X is still \mathcal{D} -dimensional, though. Further simplification is achieved by taking the McKean-Vlasov limit, which is also denoted as the many-demes limit or the mean field approximation. To take that limit, assume a finite set of demes $\mathcal{D} = \{1, \dots, D\}$, for some $D \in \mathbb{N}$, with uniform migration between demes. If exchangeability of the starting value of X across demes is assumed, then the uniform migration leads to an averaging of the immigration terms. In the limit of $D \rightarrow \infty$, this average converges to the expected value. Thus, for infinitely many demes, the altruist frequency in each deme can be described by the same one-dimensional process. These ideas are formalized in Section 4.1.5. In Proposition 4.1.13, the result is shown in a general setting and in Lemma 4.1.14, applicability to the case of altruistic defense traits is confirmed. The process describing the evolution of the altruistic defense allele is thereby reduced to one dimension. Informally speaking, these results show that if the population is structured in many demes, then the influence of any single deme is reduced and the effects of migration are averaged over the whole population. Thereby, under appropriate starting conditions, the altruist frequency in any deme can be approximated by the average altruist frequency across all demes. The longterm behavior of the resulting process is investigated in Section 4.1.6. Theorem 3.1.3 provides a simple condition that determines whether the altruistic defense allele goes to fixation or extinction. The critical case of coexistence between

altruists and cheaters is covered as well. The proof is given in Section 4.1.6.3. Theorem 3.1.3 states that in the setting of large populations structured in many demes, where altruists can be observed in each deme in the beginning, then altruists will take over the population if $\alpha < \beta$, i.e., if their “cost” is lower than their “benefit” to the host population. If this is not the case, then altruists will go extinct. Furthermore, the scenario of coexistence between altruists and cheaters would only occur in the critical case of $\alpha = \beta$, and is thus unlikely to be observed in natural populations.

Theorem 3.1.3. *Let $\alpha, \beta, \kappa \in (0, \infty)$, let $a \in (1, \infty)$, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, \infty)})$ be a filtered probability space, let $W: [0, \infty) \times \Omega \rightarrow \mathbb{R}$ be a standard $(\mathcal{F}_t)_{t \in [0, \infty)}$ Brownian motion with continuous sample paths, and let $Z_0: \Omega \rightarrow [0, 1]$ be an $\mathcal{F}_0/\mathcal{B}([0, 1])$ -measurable mapping. Then the SDE*

$$\begin{aligned} dZ_t = & \kappa(a - Z_t) \left((a - Z_t) \mathbb{E} \left[\frac{1}{a - Z_t} \right] - 1 \right) dt - \alpha Z_t (1 - Z_t) dt \\ & + \sqrt{\beta(a - Z_t)Z_t(1 - Z_t)} dW_t \end{aligned} \quad (3.4)$$

has a unique solution. Furthermore, if $\mathbb{E}[Z_0] = 1$, then $\mathbb{P}[Z_t = 1 \text{ for all } t \in [0, \infty)] = 1$, if $\mathbb{E}[Z_0] = 0$, then $\mathbb{P}[Z_t = 0 \text{ for all } t \in [0, \infty)] = 1$ and if $\mathbb{E}[Z_0] \in (0, 1)$, then

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}[|Z_t - 0|] &= 0, \text{ if } \alpha > \beta, \\ \lim_{t \rightarrow \infty} \mathbb{E}[|Z_t - 1|] &= 0, \text{ if } \alpha < \beta, \\ Z_t &\xrightarrow{t \rightarrow \infty} \int m(z) dz, \text{ if } \alpha = \beta, \end{aligned} \quad (3.5)$$

where $m(z) = \frac{1}{c} z^{\frac{2\kappa}{\beta}(a\theta-1)-1} (1-z)^{\frac{2\kappa}{\beta}(1-\theta(a-1))-1} (a-z)^{\frac{2\alpha}{\beta}-1}$ for $z \in (0, 1)$, where $c \in (0, \infty)$ is a normalizing constant and where $\theta = \mathbb{E}[\frac{1}{a-Z_0}]$.

Thus, if the altruistic defense allele starts in positive frequency in each deme in an infinite dimensional space, then it will go to fixation if $\alpha < \beta$ or it will go extinct if $\alpha > \beta$. If $\alpha = \beta$, then there will be coexistence of altruists and cheaters with the specified equilibrium distribution of the altruist frequency.

The previous statements assume that the altruistic defense allele is already present in positive frequency across infinitely many demes. The case of invasion into a population of non-altruistic individuals is thereby not covered by these results. This gap is closed in Section 4.1.7. There, the altruistic defense allele is assumed to have arisen in small frequency in a single deme, possibly due to a mutation event. Then, Proposition 4.1.17 shows that the survival probability of an invading defense allele in an infinite dimensional space is positive, if and only if $\alpha < \beta$. Thus, for large populations structured in many demes, a newly arisen

altruistic defense allele has a chance of establishing in the host population only if its “cost” is not too large ($\alpha < \beta$). Otherwise, it will surely go extinct.

3.2 Simulations on the evolution of altruistic defense

3.2.1 Small populations

To assess the robustness of the results from the theoretical analysis in a setting with small populations living in finitely many demes, several simulation series were performed. The parameter values of all simulation series are shown in Table 2.2. For each parameter configuration, 100 simulation runs were performed. Data was analyzed using *R* (R Core Team, 2015).

In *simulation series α* , the effect of finite population sizes on the theoretical prediction was tested for different values of α and b (where $b = \beta/\beta_H$). Results are shown in Figure 3.2, where each point gives the mean value of the 100 simulation runs. All final frequencies were either zero or one, i.e., extinction or fixation of altruists was observed in each run. Bars denote Wilson confidence intervals on the 5% level and the curves are obtained from logistic regression. None of the curves agree with the theoretical result that predicts extinction of altruists for $\alpha > b$ and fixation of altruists for $\alpha < b$ (note that $\beta_H = 1$ is fixed in *simulation series α* and thus, $b = \beta$). Qualitatively, the trend can be seen, but α and b have to be sufficiently far apart for fixation or extinction of altruists to be stable across different simulation runs with the same setting. For $\alpha = 1$, extinction of altruists was observed in all 100 simulations for each of the values of b . For $\alpha = 0.5$ and $\alpha = 0.4$, there was no significant difference between the results for the different values of b (Bonferroni-Holm corrected p-values of 0.37 and 0.10, respectively). For all other values of α , there was significant difference between the results for the different values of b (all Bonferroni-Holm corrected p-values below 0.0002). Thus, for the selective disadvantage of altruists to be large enough, to overcome even small benefit effects of $b \leq 0.05$, α has to be chosen sufficiently large.

Effects of randomness in host reproduction introduced by β_H were investigated in *simulation series β_H* (see Figure 3.3). Results were derived analogously to those of *simulation series α* , above. Note that the results for $b = 0.7$ were excluded for the logistic regression. There is no significant difference between the results for the different values of β_H for $b = 0.6$ (Bonferroni-Holm corrected p-value of 0.11). For all other values of b , there was a significant difference (maximum corrected p-value of 2E-6). In the simulations with $b = 0.7$, the parasite population went extinct in 94, 98, 96, and 82 runs out of the 100 runs for $\beta_H = 0$, $\beta_H = 0.1$, $\beta_H = 1$, and $\beta_H = 10$, respectively. For all other values of b , the

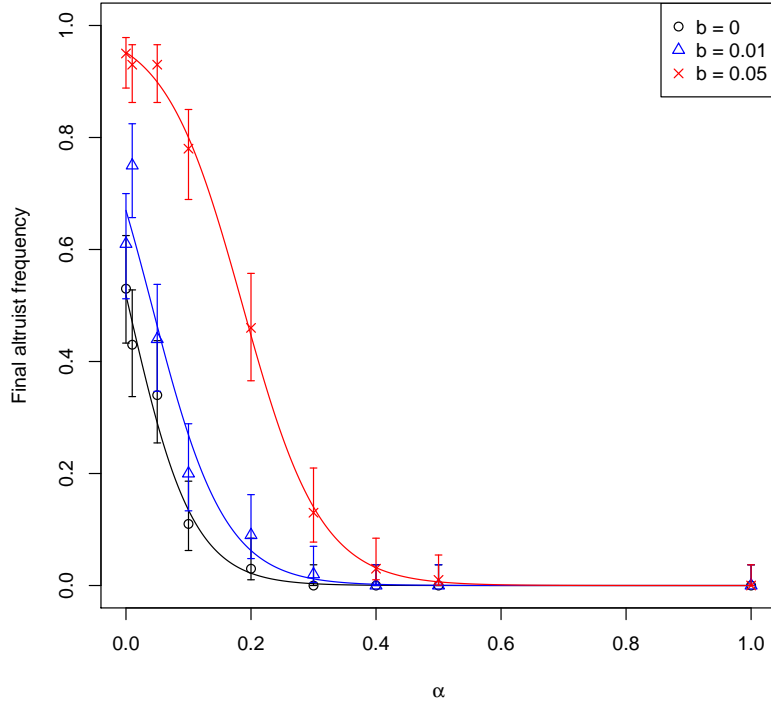


Figure 3.2: Results for *simulation series* α . Each point shows the average final altruist frequency across 100 runs. See text for details.

parasite population never went extinct in any of the simulations.

Simulation series κ_H investigated the effect of host migration rates. Results are shown in Figure 3.4. They were derived analogously to those of *simulation series* α , above. Among the simulations with $\kappa_H = 0.01$, there were eight and five runs out of 100 total, for $b = 0.2$ and $b = 0.3$, respectively, where neither extinction nor fixation of altruists had occurred by the end of the simulation. This means there were both altruists and cheaters present in the host population and final altruist frequency was strictly between zero and one. These runs were excluded from the computation of confidence intervals, logistic regression, and significance. For $b = 0$ and $b = 0.1$, all simulations led to extinction of altruists, while for $b = 0.4$ and $b = 0.5$, all simulations showed fixation of altruists. Intermediate values of b showed no significant difference between different settings of κ_H (all Bonferroni-Holm corrected p-values above 0.39).

The effect of random fluctuations in host carrying capacity was studied in *simulation series* a_f . Results were derived analogously to those of *simulation series* α , above, and are shown in Figure 3.5. For $b = 0$, all simulations led to extinction of altruists, while for $b = 0.4$ and $b = 0.5$, all simulations showed fixation of altruists. There was a significant effect of the amount of fluctuation, a_f , for $b = 0.2$ (Bonferroni-Holm corrected p-value of 5.6E-7). For $b = 0.1$ and $b = 0.3$, the effect of a_f was not significant (corrected p-value of 0.52 for both).

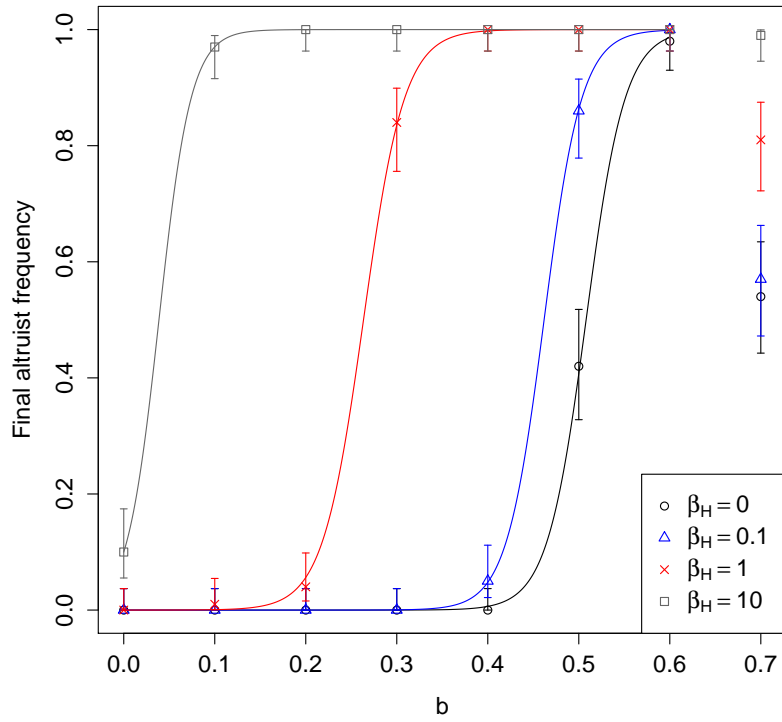


Figure 3.3: Results for *simulation series* β_H . Each point shows the average final altruist frequency across 100 runs. See text for details.

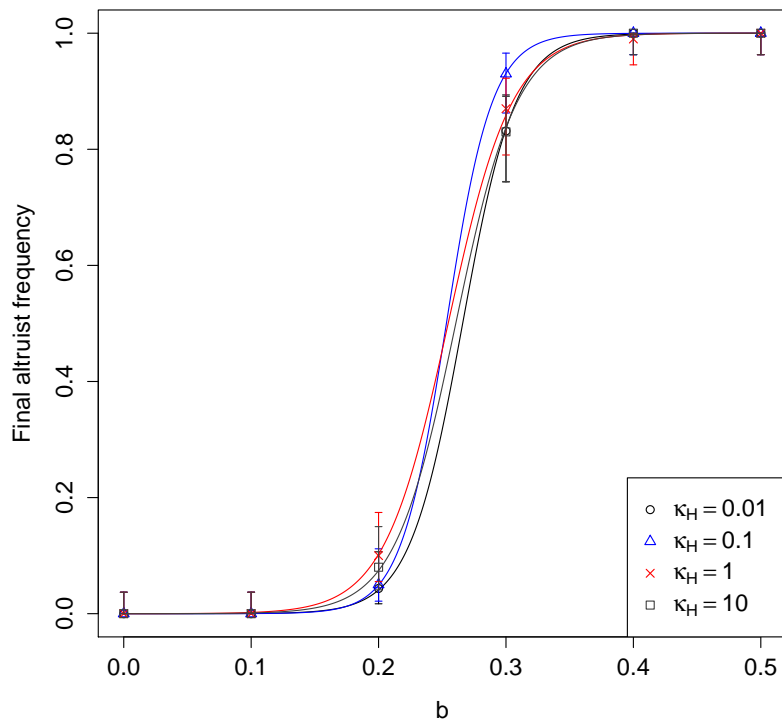


Figure 3.4: Results for *simulation series* κ_H . Each point shows the average final altruist frequency across 100 runs. See text for details.

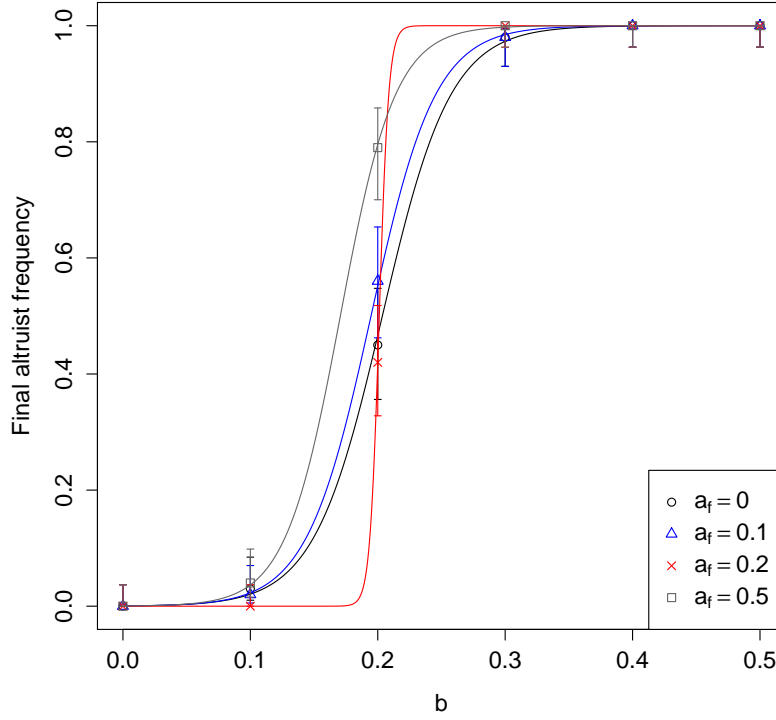


Figure 3.5: Results for *simulation series* a_f . Each point shows the average final altruist frequency across 100 runs. See text for details.

3.2.2 Large populations

Several simulation series were performed, to check how well the condition derived in the theoretical analysis could predict the outcome in a setting of infinitely large populations living in finitely many demes. The parameters of all simulation series are shown in Table 2.3. Data was analyzed using *R* (R Core Team, 2015).

First, *simulation series* ε was performed, to find a suitable cutoff value to account for the discretization (see Box 2.5). For each parameter set, 25 simulations were performed. Since in *simulation series* ε , α is set to 0.05, the predicted value is zero for $\beta < 0.05$ and one for $\beta > 0.05$ (recall that $\beta = b \cdot \beta_H$). Final altruist frequencies for all runs are shown in Figure 3.6. Points are spread out horizontally. The actual values of β are 0.01, 0.03, 0.07, and 0.09. Extinction or fixation of altruists occurred in all runs for $\varepsilon = 10^{-7}$, in 97 of the 100 runs for $\varepsilon = 10^{-6}$, in the 25 runs with $\beta = 0.01$ but never with other values of β for $\varepsilon = 10^{-8}$, and in none of the runs for $\varepsilon = 0$.

The value $\varepsilon = 10^{-7}$ was then used in *simulation series* U , N , $2D$, and T , with 100 runs per parameter set. Mean values with standard errors from these simulation series are shown in Figure 3.7 and root mean squared errors of the simulation results compared to the theoretical prediction are shown in Figure 3.8. The simulations with $\beta = \alpha$ were excluded from the

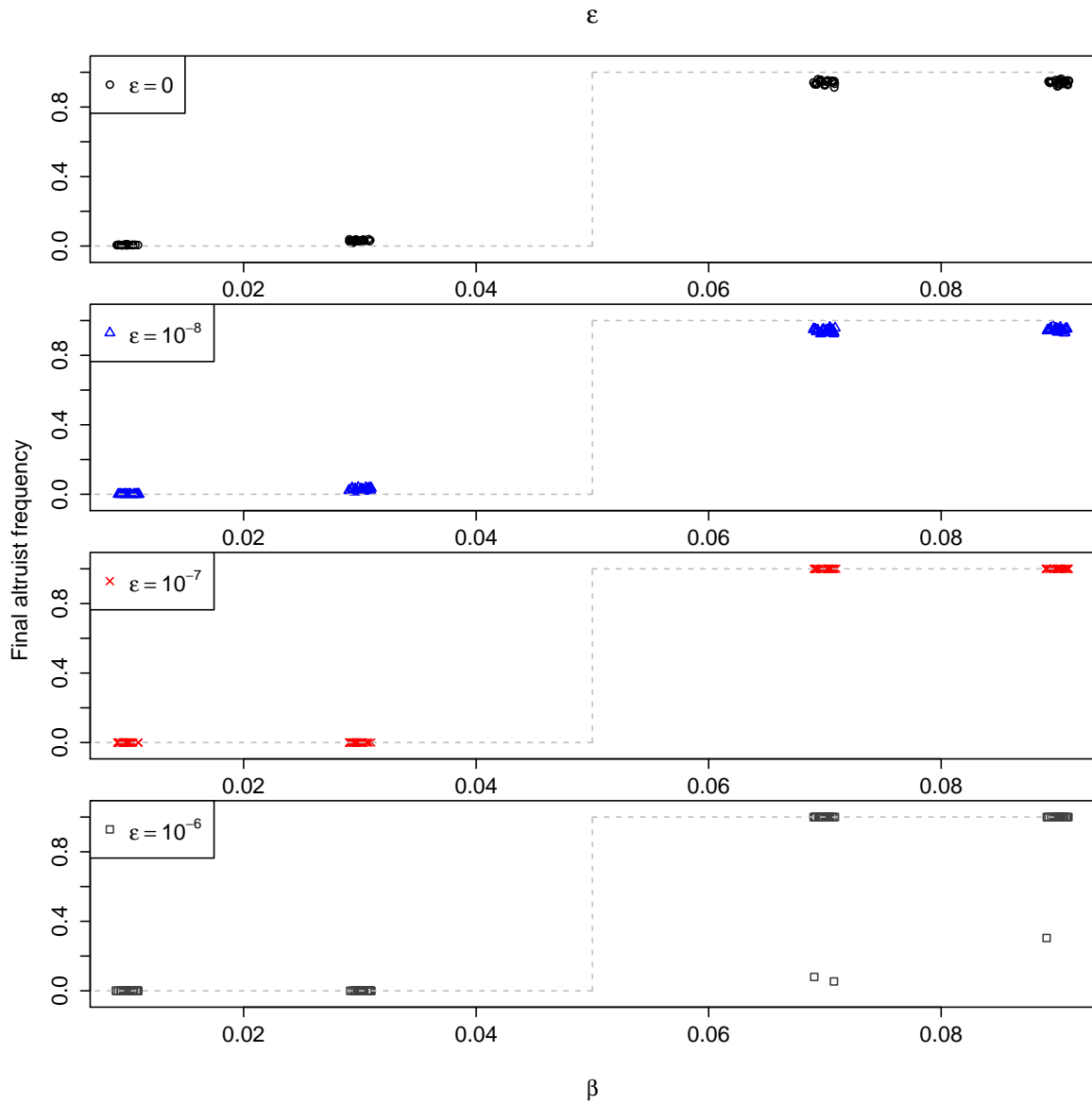


Figure 3.6: Final altruist frequencies from *simulation series* ε . Points are spread out horizontally. Dashed gray lines indicate the theoretical predictions. See text for details.

computation of root mean squared errors.

In *simulation series U*, the prediction from the theoretical results was checked for uniform migration between the demes, with different total deme numbers. For each parameter setting, 100 simulation runs were performed. For $\beta = \alpha$, there were 84 runs with $D = 500$ and one run with $D = 50$ that did not show extinction or fixation of altruists. All other runs had a final altruist frequency of either zero or one. Results are shown in Figure 3.9. Each point indicates the final altruist frequency of one simulation run. Points are spread out horizontally. The actual values of β are 0.01, 0.02, 0.03, 0.04, 0.045, 0.05, 0.055, 0.06, 0.07, 0.08, and 0.09.

In *simulation series N*, the process X was simulated under nearest-neighbor migration on a circle. For each parameter setting, 100 simulation runs were performed. For $D = 50$, 125 of the 500 runs with $0.04 \leq \beta \leq 0.06$ showed coexistence of altruists and cheaters, while for $D = 500$, coexistence was observed in 425 of the 700 runs with $0.04 \leq \beta \leq 0.08$. All other runs of the simulation series showed fixation or extinction of altruists. For $\alpha = \beta$, a final altruist frequency of either zero or one is observed in 100, 45, and 0 runs (out of a total of 100 runs) for $D = 5$, $D = 50$, and $D = 500$, respectively. Results are shown in Figure 3.10. Each point indicates the final altruist frequency of one simulation run. Points are spread out horizontally. The actual values of β are 0.01, 0.02, 0.03, 0.04, 0.045, 0.05, 0.055, 0.06, 0.07, 0.08, and 0.09.

In *simulation series 2D*, nearest-neighbor migration in two dimensions was assumed, i.e., each deme sent migrants to the four immediate neighbors on a two-dimensional grid (with a torus structure to avoid boundary effects). For each parameter setting, 100 simulation runs were performed. In none of the simulations, fixation or extinction of altruists occurred. Results are shown in Figure 3.11. Each point indicates the final altruist frequency of one simulation run. Points are spread out horizontally. The actual values of β are 0.02, 0.04, 0.06, 0.08, 0.09, 0.1, 0.11, 0.12, 0.14, 0.16, and 0.18.

The process X was simulated with migration along edges of a binary tree in *simulation series T*. Demes were represented as nodes of a binary tree and each node exchanged migrants with its parent and its two children. The root sent migrants only to its two children. Neighboring leaves also exchanged migrants. For each parameter setting, 100 simulation runs were performed. In all simulations, the final altruist frequency was strictly between zero and one. Results are shown in Figure 3.12. Each point indicates the final altruist frequency of one simulation run. Points are spread out horizontally. The actual values of β are 0.01, 0.02, 0.03, 0.04, 0.045, 0.05, 0.055, 0.06, 0.07, 0.08, and 0.09.

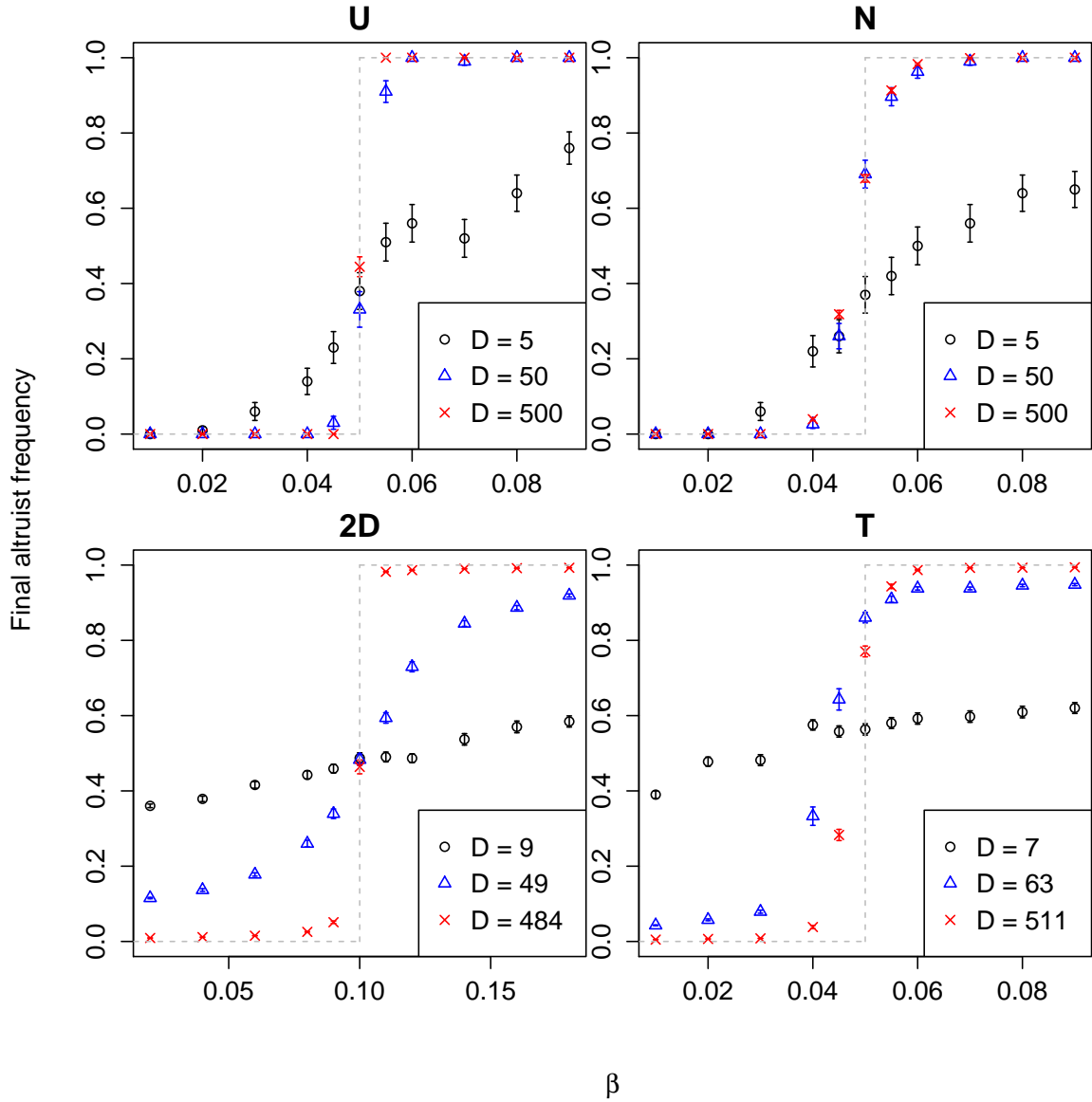


Figure 3.7: Mean values and standard errors from 100 simulation runs with uniform migration (U), nearest-neighbor migration (N), two-dimensional nearest-neighbor migration (2D), and migration along edges of a binary tree (T). Dashed gray lines indicate the theoretical predictions.

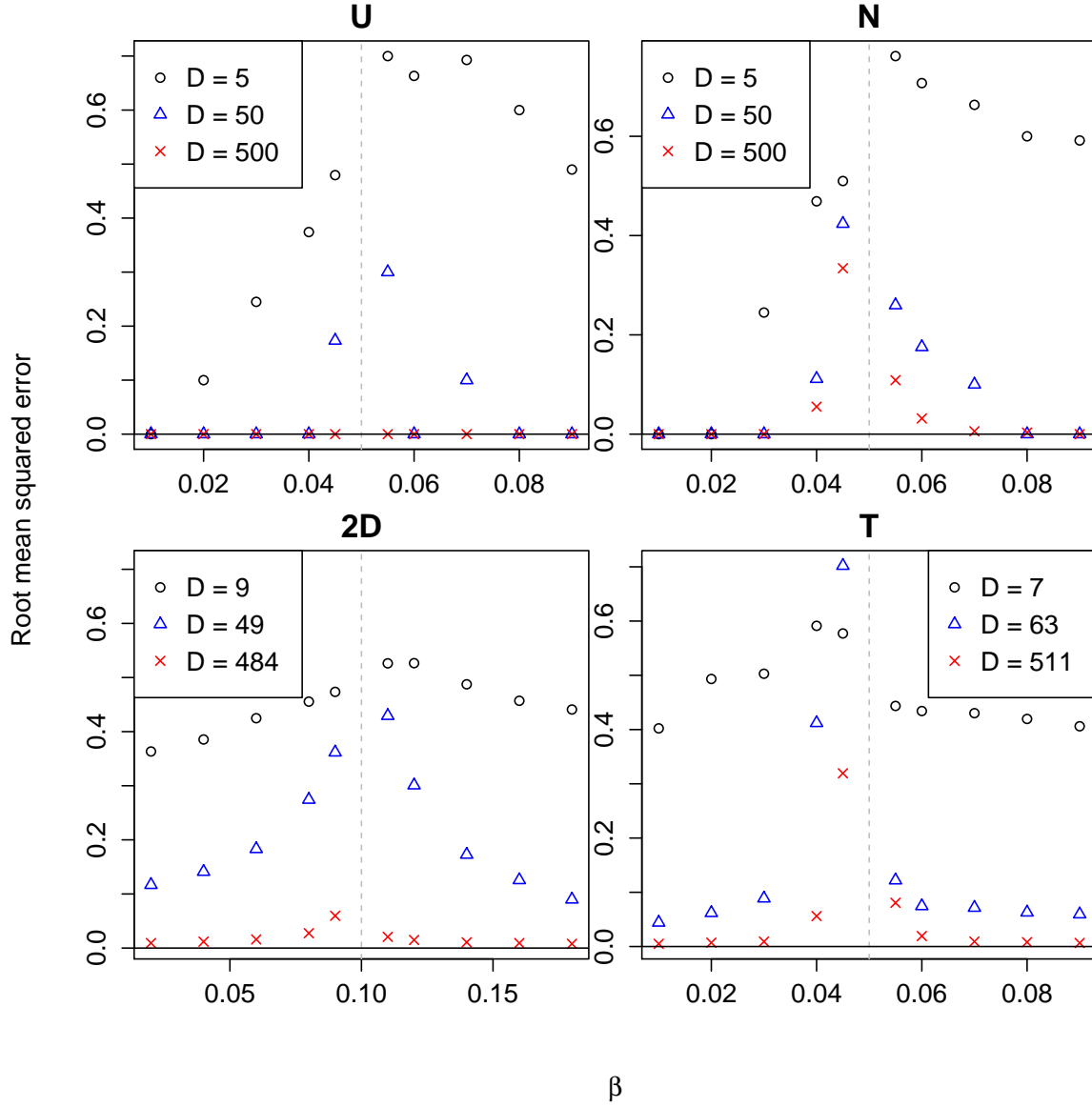


Figure 3.8: Root mean squared errors from 100 simulation runs with uniform migration (U), nearest-neighbor migration (N), two-dimensional nearest-neighbor migration (2D), and migration along edges of a binary tree (T). Dashed gray lines indicate the respective values of α .

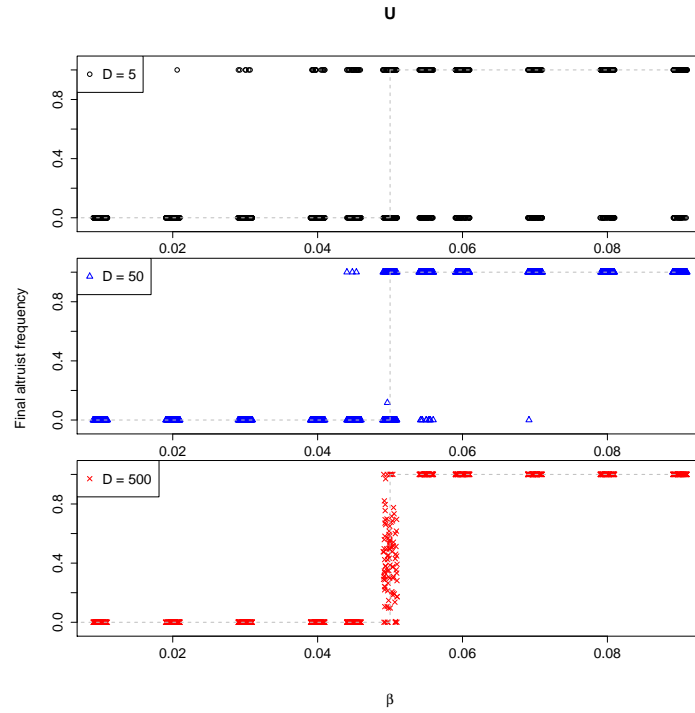


Figure 3.9: Final altruist frequencies from *simulation series U*. Points are spread out horizontally. Dashed gray lines indicate the theoretical predictions. See text for details.

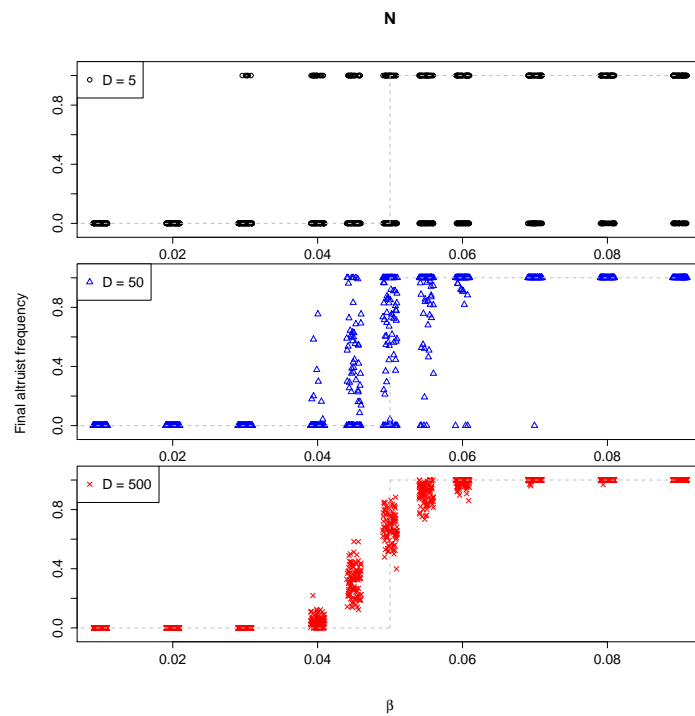


Figure 3.10: Final altruist frequencies from *simulation series N*. Points are spread out horizontally. Dashed gray lines indicate the theoretical predictions. See text for details.

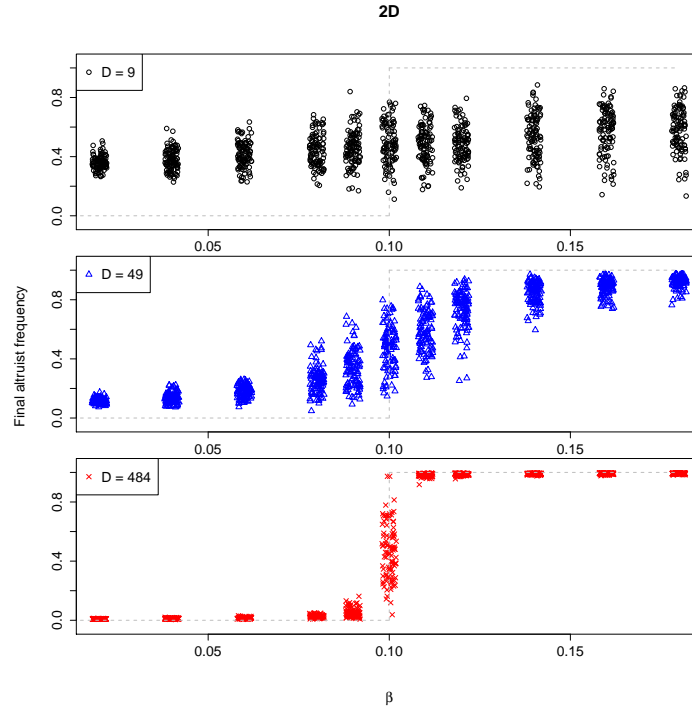


Figure 3.11: Final altruist frequencies from *simulation series 2D*. Points are spread out horizontally. Dashed gray lines indicate the theoretical predictions. See text for details.

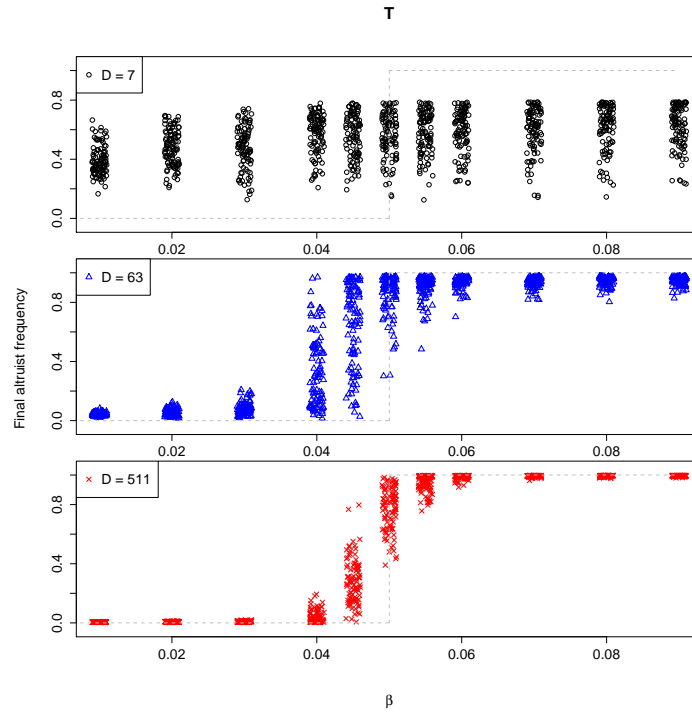


Figure 3.12: Final altruist frequencies from *simulation series T*. Points are spread out horizontally. Dashed gray lines indicate the theoretical predictions. See text for details.

3.3 Evolution of slave rebellion in ants

In this section, the results of Metzler et al. (2015) are described. The data was analyzed using *R* (R Core Team, 2015).

Prior simulations

The *Prior simulations* were used in an Approximate Bayesian Computation (ABC) analysis, to find parameter sets that best suited the field observations. From the field data of Pamminger et al. (2014), which comprises two locations—New York State (NY) and West Virginia (WV)—14 summary statistics were computed. The values for these summary statistics are provided in Table 2.7. From the 10 000 runs of the *Prior simulations*, the 5% were kept that resulted in summary statistics closest to the observed values (under Euclidean distance), separately for both locations. For some of the model parameters, this resulted in a narrowed down range, compared to the prior distributions. Notably, the distance in which a host nest can build satellite nests was below 1 m, for all of the selected parameter combinations. However, for other parameters, such as the maximum numbers of host and slavemaker workers per nest, the data was not informative, such that the posterior distributions closely resembled the prior distributions.

Posterior simulations

For each of the 500 parameter sets per location chosen in the ABC analysis of the *Prior simulations*, ten runs were performed. The rebel allele frequency in the beginning of the simulation and in immigrating host queens and males was set to 0.01. Simulations with a final host nest density below 0.1 per m² were excluded from further analyses. This led to a remaining number of 4063 runs for the NY *Posterior simulations* and 3847 for the WV *Posterior simulations*. Among the remaining simulations, there were 1780 for NY and 1545 for WV with the cost parameter of the rebellion allele above 10%. For these runs, the final rebel allele frequency never exceeded the initial value of 0.01. In a few runs with cost parameter below 1%, the final rebel allele frequency was above that initial value. For NY and WV combined, there was a total of 4212 simulations with final host nest density above 0.1 per m² and cost parameter between 0.5% and 10%. Among these, four NY simulations and seven WV simulations showed a final rebel allele frequency higher than 0.01. Furthermore, there were 27 runs for NY and 34 runs for WV, which showed a long-term increased rebel nest frequency over the last 1000 years of the simulation.

Posterior simulations with fixed cost

In this simulation series, the same 1000 parameter sets as of the *Posterior simulations* were used, but the cost of the rebel allele was set to 0.01 or 0.05. For both cost settings, three runs were performed per parameter set. There were 26 parameter combinations with a cost of 0.01 and 11 with a cost of 0.05 that showed a long-term rebel nest frequency above 0.05 in at least one of the three simulation runs.

Posterior simulations without slavemakers

For each of the parameter sets of the *Posterior simulations*, three additional runs were performed, in which slavemakers were completely absent. Among the 3000 simulations, there were 1515 where the cost parameter was between 0.5% and 10% and the final host nest density was above 0.1 per m². In all of these runs, the final rebel allele frequency was below the initial value of 0.01. In six of the remaining 1485 simulations, the final rebel allele frequency was slightly above the initial value, with a maximum value of 0.016. The cost parameter in these six simulations was never larger than 0.21%. Furthermore, for the NY data, there were 102 parameter combinations, in which either the mean final rebel allele frequency of the three simulations from this series or the mean final rebel allele frequency from the first three runs of the corresponding *Posterior simulations* was above 10^{-5} . For the WV data, there were 114 such parameter combinations. In both locations, the mean was higher for the simulations with slavemakers in the majority of the cases ($^{64}_{102}$ for NY and $^{71}_{114}$ for WV). This over-representation was significant in both locations ($p = 0.013$ for NY and $p = 0.011$ for WV, two-sided test based on binomial distribution).

Simulations with potential for rebellion

For each of the 61 parameter sets identified in the *Posterior simulations* that showed a long-term increased rebel nest frequency, additional 20 simulations were performed. Among these 1220 runs, there were 81 that again showed an average rebel nest frequency over 0.05 in the last 1000 simulated years. In a large number of these, a pattern of extinction and recolonization can be observed. Whenever the amount of slavemaker nests on the grid reaches a critical number, all host nests die shortly afterwards. Due to a lack of hosts, the slavemakers then quickly vanish as well. An illustration of this pattern of frequent extinction and recolonization is provided in Figure 3.13, where nest numbers of three simulations from a single parameter combination with cost 0.01 are shown. In rare cases, a slow increase of the rebel nest frequency over time could be observed, but only if the cost parameter was very low. An

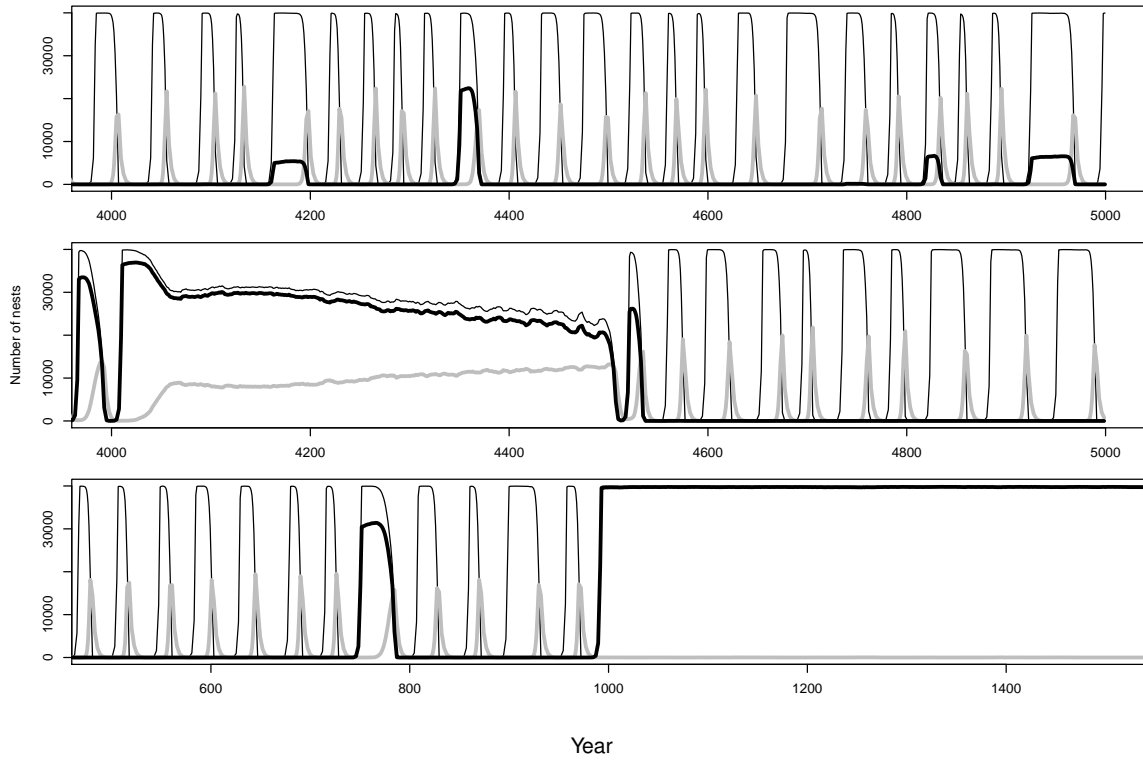


Figure 3.13: Numbers of host nests (black), rebel host nests (bold black), and slavemaker nests (gray) over time in three simulations of the same parameter combination with cost parameter of 0.01. The first two plots show the last 1000 years of the simulations, while the third one shows years 500 to 1500 of a simulation that had constant nest numbers from year 1000 onwards. Figure taken from Metzler et al. (2015).

example with cost 0.000 59 is shown in Figure 3.14.

Simulations with potential for rebellion with fixed cost

There were 37 parameter sets in the *Posterior simulations with fixed cost* for which a long-term rebel nest frequency above 0.05 during the last 1000 simulated years was found in at least one of the three simulation runs. For each of these parameter sets, 40 additional simulations were performed and checked for the criteria that (a) during the last 1000 years of the simulation, the average host nest density was larger than 0.1 per m² and (b) during the last 1000 years of the simulation, the rebel allele frequency among all produced host queens and males was larger than 0.01. For 6 of the 26 parameter sets with cost 0.01, and for 4 of the 11 parameter sets with cost 0.05, none of the 40 simulations met both criteria (a) and (b) combined. The majority of the 37 parameter sets fulfilled (a) and (b) together in at most 10 of the 40 simulations. Only three parameter sets fulfilled (a) and (b) in more simulations. Two parameter combinations with cost 0.01 met both (a) and (b) in 16 and 17 out of the 40

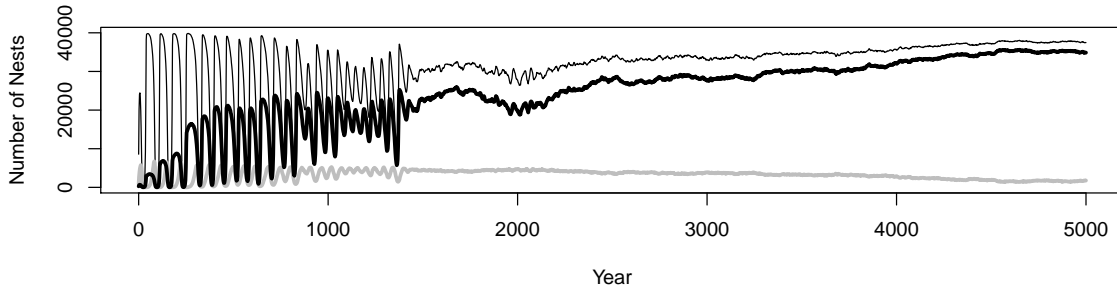


Figure 3.14: Numbers of host nests (black), rebel host nests (bold black), and slavemaker nests (gray) over time in a simulation with cost parameter of 0.000 59 and an average rebel nest frequency in the last 1000 years of 0.93. Figure taken from Metzler et al. (2015).

runs, respectively, while one parameter combination with cost 0.05 fulfilled both criteria in 16 simulations. For one of the parameter sets with cost 0.01 the nest numbers over time of three simulations are shown in Figure 3.13. Although these simulations stem from the same parameter set, the plots in Figure 3.13 show different outcomes. This illustrates the important role that stochasticity plays in the simulation model. The pattern of frequent extinction and recolonization, as described in the previous section for *Simulations with potential for rebellion*, occurred in all simulations that showed a long-term increased rebel nest frequency. When the rebel frequency was high during the recolonization of an empty grid, this could protect the host against invading slavemakers and thereby lead to a prolonged period of host prevalence. Thus, host nests could be more productive in the years following a recolonization with an increased rebel frequency. To test this effect, the rebel allele frequency among all produced host queens and males during the last 1000 years was averaged over all 40 simulation runs for all 37 parameter sets. For the 26 parameter combinations with cost 0.01, there were 16 with an average rebel allele frequency among produced sexuals during the last 1000 years over 0.01, with a maximum value of 0.043. Among these, 13 had an average host nest density above 0.1 per m^2 during the last 1000 years. For the 11 parameter combinations with cost 0.05, there were 5 with an average rebel allele frequency among produced sexuals during the last 1000 years over 0.01, with a maximum value of 0.027. Among these, 3 had an average host nest density above 0.1 per m^2 during the last 1000 years. See Figure 3.15 for plots of the average rebel allele frequencies. These 16 parameter combinations with cost 0.01 or 0.05 that showed an increase in rebel frequency among produced sexuals and a sufficiently large nest density, were then compared to a random sample of the 1000 parameter sets of the *Posterior simulations with fixed cost*. The parameters μ_{xp} and λ_h were the only ones showing a significant difference on the 5% level, with p -values of 0.000 33 and 0.0062, respectively

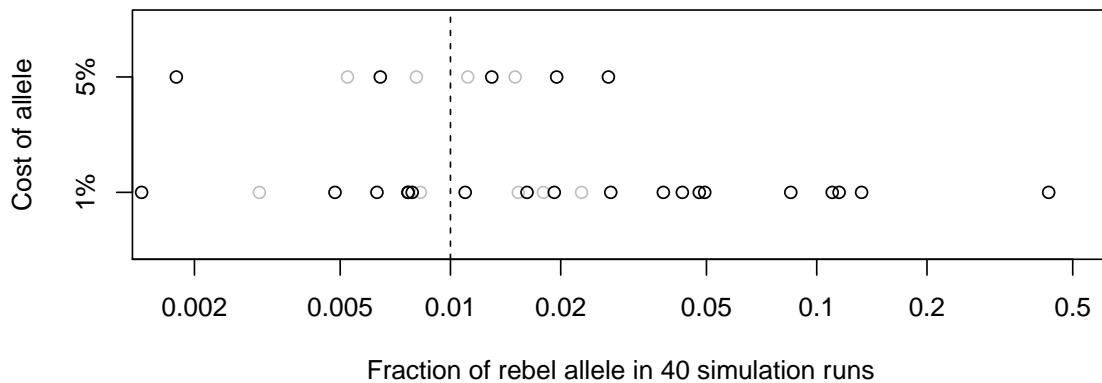


Figure 3.15: Rebel allele frequencies among all host queens and males produced in the last 1000 years. Each dot shows the value for all 40 simulations of one parameter combination. A dot is gray, if the average host nest density in the last 1000 years across all 40 simulations was below 0.1 per m^2 . The dashed line shows the initial value of 0.01. Results for parameter combinations with an average rebel allele frequency below 0.001 are not shown. Figure taken from Metzler et al. (2015).

(Wilcoxon rank test with Bonferroni-Holm correction; see Figure 3.16 and Table 2.4). The former parameter describes the death rate of slavemaker queens returning from their nuptial flights and a low value is advantageous for the rebel allele. The latter controls how many new host workers are produced by host nests, depending on the current worker number and a large value is advantageous for the rebel allele. There may be more parameters that are relevant in the spread of the rebel allele that could not be detected by the amount of simulations, because their effects are too weak or their interactions too complicated. Furthermore, only the range of posterior distributions was covered by the analysis. This means that some possibly relevant parameters were already rather restricted, such as the host nest branching distance.

Metapopulation simulations

For the three parameter sets of the *Simulations with potential for rebellion with fixed cost* that fulfilled the two criteria that (a) during the last 1000 years of the simulation, the average host nest density was larger than 0.1 per m^2 and (b) during the last 1000 years of the simulation, the rebel allele frequency among all produced host queens and males was larger than 0.01 in more than a quarter of the 40 simulation runs, additional ten simulations were performed in a

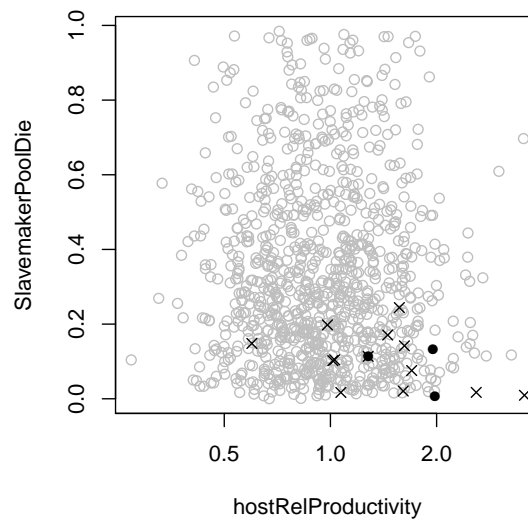


Figure 3.16: Parameter values of μ_{xp} (“SlavemakerPoolDie”) and λ_h (“hostRelProductivity”) from all posterior samples of *Posterior simulations* for NY and WV combined. Black crosses show the 13 parameter values of parameter sets with cost 0.01 chosen from the *Simulations with potential for rebellion with fixed cost* with an increased rebel allele frequency among all newly produced host queens and males and a host nest density of at least 0.1 per m², both in the last 1000 years. Black dots show the corresponding five values from parameter sets with cost 0.05. Figure taken from Metzler et al. (2015).

metapopulation scenario. In each simulation, 20 demes were connected and the within-deme and between-deme components of Queller's variant of the Price equation were computed every ten years. For the parameter set with cost 0.05, the rebel allele never stabilized at a value significantly higher than the initial value of 0.01 in any of the ten simulations. For one of the two parameter sets with cost 0.01, the rebel allele nearly reached fixation in six of the ten simulations, and for the other parameter set, the rebel allele frequency went up to over 90% in three of the ten simulations. At the time points when the rebel allele frequency increased in value, the between-deme component of Queller's variant of the Price equation was clearly enlarged. On the other hand, the within-deme component stayed close to zero with only a few exceptions of short time spans. The rebel allele frequencies and the components of Queller's variant of the Price equation are shown in Figure 3.17 for all nine simulations showing an increased rebel frequency.

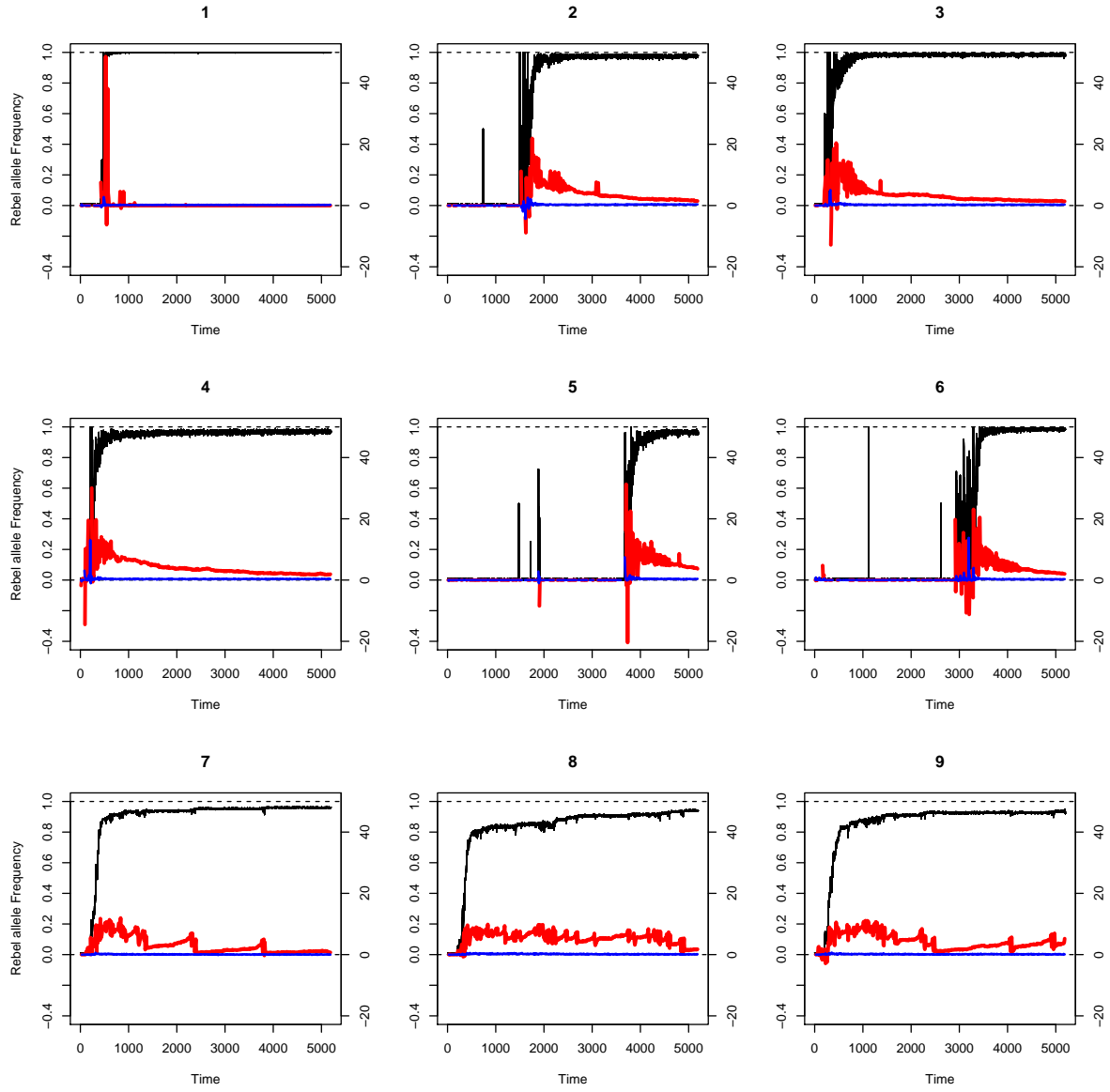


Figure 3.17: Rebel allele frequencies (black, left axis) as well as between-deme component (red, right axis) and within-deme component (blue, right axis) of Queller's variant of the Price equation over time from two different parameter sets with cost 0.01. Plots 1 to 6 and plots 7 to 9 correspond to one parameter set, respectively. Figure taken from Metzler et al. (2015).

Chapter 4

Proofs

4.1 Mathematical analysis and asymptotics

4.1.1 Notation

Throughout Section 4.1, the following notation will be used. Define $[0, \infty] := [0, \infty) \cup \{\infty\}$. Assume the conventions that $0^0 = 1$, $0 \cdot \infty = 0$, that for any $x \in (0, \infty)$ it holds that $\frac{x}{\infty} = 0$ and $\frac{x}{0} = \infty$, and that zero times an undefined expression is set to zero. For all $x, y \in \mathbb{R}$ define $x^+ := \max\{x, 0\}$, $\text{sgn}(x) := \mathbb{1}_{x>0} - \mathbb{1}_{x<0}$, and $x \wedge y := \min\{x, y\}$. Let $\sup(\emptyset) := -\infty$ and $\inf(\emptyset) := \infty$. For a topological space (E, \mathcal{E}) denote by $\mathcal{B}(E)$ the Borel sigma-algebra of (E, \mathcal{E}) . For every countable set \mathcal{D} and every $\sigma = (\sigma_i)_{i \in \mathcal{D}} \in (0, \infty)^\mathcal{D}$ define a function $\|\cdot\|_\sigma: \mathbb{R}^\mathcal{D} \rightarrow [0, \infty]$ by $\mathbb{R}^\mathcal{D} \ni z = (z_i)_{i \in \mathcal{D}} \mapsto \|z\|_\sigma := \sum_{i \in \mathcal{D}} \sigma_i |z_i|$ and define $l_\sigma^1 := \{z \in \mathbb{R}^\mathcal{D} : \|z\|_\sigma < \infty\}$.

4.1.2 Diffusion approximation

In this section, the diffusion approximation of the Markov chain given in Section 2.1.1 is derived.

Let \mathcal{D} be a finite or countable set, that denotes the set of demes and let $m \in [0, \infty)^{\mathcal{D} \times \mathcal{D}}$ such that for every $i \in \mathcal{D}$ it holds that $\sum_{k \in \mathcal{D}} m(k, i) = \sum_{k \in \mathcal{D}} m(i, k) = 1$. Furthermore, let $\lambda, K, \delta, \beta_H, \nu, \gamma, \eta, \rho, \beta_P, \alpha, \kappa_H, \kappa_P \in (0, \infty)$, such that $\rho < \eta$ and let $\sigma \in [0, \infty)^\mathcal{D}$ such that $\sum_{i \in \mathcal{D}} \sigma_i < \infty$. Define $E_2 := l_\sigma^1 \cap [0, \infty)^\mathcal{D}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For $i \in \mathcal{D}$, $N \in \mathbb{N}$, and $t \in [0, \infty)$, denote by $\overline{A}_t^N(i)$, $\overline{C}_t^N(i)$, and $\overline{P}_t^N(i)$, the numbers of altruists, cheaters, and parasites at time t , in deme i , respectively. The diffusion approximation will be performed with unscaled parameters and the scaling will be introduced for the resulting diffusions to arrive at the appropriate model. This means that β_H is replaced by

$\tilde{\beta}_H N$, β_P by $\tilde{\beta}_P N$, $\frac{\kappa_H}{N}$ by $\tilde{\kappa}_H$, $\frac{\kappa_P}{N}$ by $\tilde{\kappa}_P$, and $\frac{\alpha}{N}$ by $\tilde{\alpha}$. For $i \in \mathcal{D}$, let $a_i, c_i, p_i \in \mathbb{N}_0$. When $(\bar{A}_t^N(i), \bar{C}_t^N(i), \bar{P}_t^N(i)) = (a_i, c_i, p_i)$, $i \in \mathcal{D}$, assume these processes to have transition rates given by (2.1) with parameter substitutions as stated above, i.e.,

$$\begin{aligned}
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k + \mathbb{1}_{k=i}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \left(\tilde{\beta}_H N + \lambda \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k - \mathbb{1}_{k=i}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \left(\tilde{\beta}_H N + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} + \tilde{\alpha} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k + \mathbb{1}_{k=i}, p_k)_{k \in \mathcal{D}}: & c_i \left(\tilde{\beta}_H N + \lambda \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k - \mathbb{1}_{k=i}, p_k)_{k \in \mathcal{D}}: & c_i \left(\tilde{\beta}_H N + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k + \mathbb{1}_{k=i})_{k \in \mathcal{D}}: & p_i \left(\tilde{\beta}_P N + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k - \mathbb{1}_{k=i})_{k \in \mathcal{D}}: & p_i \left(\tilde{\beta}_P N + \nu + \gamma \frac{p_i}{N} \right) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}, c_k, p_k)_{k \in \mathcal{D}}: & a_i \tilde{\kappa}_H m(j, i) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}, p_k)_{k \in \mathcal{D}}: & c_i \tilde{\kappa}_H m(j, i) \\
 (a_k, c_k, p_k)_{k \in \mathcal{D}} &\rightarrow (a_k, c_k, p_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j})_{k \in \mathcal{D}}: & p_i \tilde{\kappa}_P m(j, i).
 \end{aligned} \tag{4.1}$$

For $i \in \mathcal{D}$, $N \in \mathbb{N}$, and $t \in [0, \infty)$, denote by $\hat{A}_t^N(i)$, $\hat{C}_t^N(i)$, and $\hat{P}_t^N(i)$, the respective population sizes of altruists, cheaters and parasites, when counted in units of N individuals, i.e., $\hat{A}_t^N(i) = \frac{\bar{A}_t^N(i)}{N}$, $\hat{C}_t^N(i) = \frac{\bar{C}_t^N(i)}{N}$, and $\hat{P}_t^N(i) = \frac{\bar{P}_t^N(i)}{N}$. Define for $N \in \mathbb{N}$, $t \in [0, \infty)$, the notation $Z_t^N := \left(\hat{A}_t^N(k), \hat{C}_t^N(k), \hat{P}_t^N(k) \right)_{k \in \mathcal{D}}$.

Denote by $C(E_2^3, \mathbb{R})$ the set of real-valued continuous functions on E_2^3 and by $C_b(E_2^3, \mathbb{R})$ ($C_b^2(E_2^3, \mathbb{R})$) the set of real-valued bounded, continuous functions (that are twice continuously differentiable, with bounded first and second order partial derivatives) on E_2^3 . Define $\text{Dom}(\mathcal{G}) := \{f \in C_b^2(E_2^3, \mathbb{R}) : f \text{ depends only on finitely many coordinates}\}$. For $N \in \mathbb{N}$, define the operator $\mathcal{G}_N : \text{Dom}(\mathcal{G}) \rightarrow C(E_2^3, \mathbb{R})$, for all $f \in \text{Dom}(\mathcal{G})$ and all $z^N = \left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k}{N} \right)_{k \in \mathcal{D}} \in E_2^3$, such that $\mathbb{P}[Z_0^N = z^N] = 1$, by

$$\mathcal{G}_N f(z^N) = \lim_{t \rightarrow 0} \frac{\mathbb{E}[f(Z_t^N) - f(z^N)]}{t}. \tag{4.2}$$

For all $N \in \mathbb{N}$, for all $f \in \text{Dom}(\mathcal{G})$, and for all $z^N = \left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k}{N} \right)_{k \in \mathcal{D}} \in E_2^3$, such that

$\mathbb{P}[Z_0^N = z^N] = 1$, it holds that

$$\begin{aligned}
 \mathcal{G}_N f(z^N) &= \lim_{t \rightarrow 0} \frac{1}{t} \left[t \sum_{i \in \mathcal{D}} \left(a_i \left(\tilde{\beta}_H N + \lambda \right) f\left(\left(\frac{a_k + \mathbb{1}_{k=i}}{N}, \frac{c_k}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}}\right) \right. \right. \\
 &\quad + a_i \left(\tilde{\beta}_H N + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} + \tilde{\alpha} \right) f\left(\left(\frac{a_k - \mathbb{1}_{k=i}}{N}, \frac{c_k}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad + c_i \left(\tilde{\beta}_H N + \lambda \right) f\left(\left(\frac{a_k}{N}, \frac{c_k + \mathbb{1}_{k=i}}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad + c_i \left(\tilde{\beta}_H N + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} \right) f\left(\left(\frac{a_k}{N}, \frac{c_k - \mathbb{1}_{k=i}}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad + p_i \left(\tilde{\beta}_P N + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) f\left(\left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k + \mathbb{1}_{k=i}}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad + p_i \left(\tilde{\beta}_P N + \nu + \gamma \frac{p_i}{N} \right) f\left(\left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k - \mathbb{1}_{k=i}}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad + \sum_{j \in \mathcal{D}} a_i \tilde{\kappa}_H m(j, i) f\left(\left(\frac{a_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}}{N}, \frac{c_k}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad + \sum_{j \in \mathcal{D}} c_i \tilde{\kappa}_H m(j, i) f\left(\left(\frac{a_k}{N}, \frac{c_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}}\right) \\
 &\quad \left. + \sum_{j \in \mathcal{D}} p_i \tilde{\kappa}_P m(j, i) f\left(\left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k - \mathbb{1}_{k=i} + \mathbb{1}_{k=j}}{N}\right)_{k \in \mathcal{D}}\right) \right] + f(z^N) \left(\mathbb{P}[Z_t^N = z^N] - 1 \right). \tag{4.3}
 \end{aligned}$$

For all $N \in \mathbb{N}$, for all $f \in \text{Dom}(\mathcal{G})$, and for all $z^N = \left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}} \in E_2^3$, such that $\mathbb{P}[Z_0^N = z^N] = 1$, it follows by Taylor approximation that

$$\begin{aligned}
 \mathcal{G}_N f(z^N) &= \sum_{i \in \mathcal{D}} \left\{ a_i \left(\tilde{\beta}_H N + \lambda \right) \left(\frac{d}{da_i} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{da_i^2} f(z^N) \frac{1}{N^2} \right) \right. \\
 &\quad + a_i \left(\tilde{\beta}_H N + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} + \tilde{\alpha} \right) \left(- \frac{d}{da_i} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{da_i^2} f(z^N) \frac{1}{N^2} \right) \\
 &\quad + c_i \left(\tilde{\beta}_H N + \lambda \right) \left(\frac{d}{dc_i} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dc_i^2} f(z^N) \frac{1}{N^2} \right) \\
 &\quad + c_i \left(\tilde{\beta}_H N + \frac{\lambda}{K} \frac{a_i + c_i}{N} + \delta \frac{p_i}{N} \right) \left(- \frac{d}{dc_i} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dc_i^2} f(z^N) \frac{1}{N^2} \right) \\
 &\quad + p_i \left(\tilde{\beta}_P N + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) \left(\frac{d}{dp_i} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dp_i^2} f(z^N) \frac{1}{N^2} \right) \\
 &\quad + p_i \left(\tilde{\beta}_P N + \nu + \gamma \frac{p_i}{N} \right) \left(- \frac{d}{dp_i} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dp_i^2} f(z^N) \frac{1}{N^2} \right) \\
 &\quad + \sum_{j \in \mathcal{D}} \left[a_i \tilde{\kappa}_H m(j, i) \left(- \frac{d}{da_i} f(z^N) \frac{1}{N} + \frac{d}{da_j} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{da_i da_j} f(z^N) \frac{1}{N^2} \right) \right. \\
 &\quad + c_i \tilde{\kappa}_H m(j, i) \left(- \frac{d}{dc_i} f(z^N) \frac{1}{N} + \frac{d}{dc_j} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dc_i dc_j} f(z^N) \frac{1}{N^2} \right) \\
 &\quad \left. + p_i \tilde{\kappa}_P m(j, i) \left(- \frac{d}{dp_i} f(z^N) \frac{1}{N} + \frac{d}{dp_j} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dp_i dp_j} f(z^N) \frac{1}{N^2} \right) \right] + O\left(\frac{1}{N}\right) \Big\}. \tag{4.4}
 \end{aligned}$$

First the migration terms will be simplified. Using the fact that for all $i \in \mathcal{D}$, it holds that $\sum_{k \in \mathcal{D}} m(i, k) = \sum_{k \in \mathcal{D}} m(k, i) = 1$, and rearranging terms, then for all $N \in \mathbb{N}$, for all $f \in \text{Dom}(\mathcal{G})$, and for all $z^N = \left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}} \in E_2^3$, such that $\mathbb{P}[Z_0^N = z^N] = 1$, it follows that

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{D}} \left[a_i \tilde{\kappa}_H m(j, i) \left(-\frac{d}{da_i} f(z^N) \frac{1}{N} + \frac{d}{da_j} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{da_i da_j} f(z^N) \frac{1}{N^2} \right) \right. \\
 & \quad + c_i \tilde{\kappa}_H m(j, i) \left(-\frac{d}{dc_i} f(z^N) \frac{1}{N} + \frac{d}{dc_j} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dc_i dc_j} f(z^N) \frac{1}{N^2} \right) \\
 & \quad \left. + p_i \tilde{\kappa}_P m(j, i) \left(-\frac{d}{dp_i} f(z^N) \frac{1}{N} + \frac{d}{dp_j} f(z^N) \frac{1}{N} + \frac{1}{2} \frac{d^2}{dp_i dp_j} f(z^N) \frac{1}{N^2} \right) \right] \\
 &= \sum_{i \in \mathcal{D}} \left[-\frac{a_i}{N} \tilde{\kappa}_H \frac{d}{da_i} f(z^N) + \sum_{j \in \mathcal{D}} \frac{a_j}{N} \tilde{\kappa}_H m(i, j) \frac{d}{da_i} f(z^N) - \frac{c_i}{N} \tilde{\kappa}_H \frac{d}{dc_i} f(z^N) \right. \\
 & \quad + \sum_{j \in \mathcal{D}} \frac{c_j}{N} \tilde{\kappa}_H m(i, j) \frac{d}{dc_i} f(z^N) - \frac{p_i}{N} \tilde{\kappa}_P \frac{d}{dp_i} f(z^N) + \sum_{j \in \mathcal{D}} \frac{p_j}{N} \tilde{\kappa}_P m(i, j) \frac{d}{dp_i} f(z^N) \\
 & \quad \left. + \sum_{j \in \mathcal{D}} m(j, i) O\left(\frac{1}{N}\right) \right] \\
 &= \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{D}} \left[\tilde{\kappa}_H m(i, j) \left(\frac{a_j}{N} - \frac{a_i}{N} \right) \frac{d}{da_i} f(z^N) + \tilde{\kappa}_H m(i, j) \left(\frac{c_j}{N} - \frac{c_i}{N} \right) \frac{d}{dc_i} f(z^N) \right. \\
 & \quad \left. + \tilde{\kappa}_P m(i, j) \left(\frac{p_j}{N} - \frac{p_i}{N} \right) \frac{d}{dp_i} f(z^N) \right] + \sum_{i \in \mathcal{D}} O\left(\frac{1}{N}\right). \tag{4.5}
 \end{aligned}$$

Consequently, for all $N \in \mathbb{N}$, all $f \in \text{Dom}(\mathcal{G})$, and all $z^N = \left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k}{N}\right)_{k \in \mathcal{D}} \in E_2^3$, such that $\mathbb{P}[Z_0^N = z^N] = 1$, it holds that

$$\begin{aligned}
 \mathcal{G}_N f(z^N) &= \sum_{i \in \mathcal{D}} \left\{ \frac{a_i}{N} \left(\lambda - \frac{\lambda}{K} \frac{a_i + c_i}{N} - \delta \frac{p_i}{N} - \tilde{\alpha} \right) \frac{d}{da_i} f(z^N) + \frac{a_i}{N} \tilde{\beta}_H \frac{d^2}{da_i^2} f(z^N) \right. \\
 & \quad + \frac{c_i}{N} \left(\lambda - \frac{\lambda}{K} \frac{a_i + c_i}{N} - \delta \frac{p_i}{N} \right) \frac{d}{dc_i} f(z^N) + \frac{c_i}{N} \tilde{\beta}_H \frac{d^2}{dc_i^2} f(z^N) \\
 & \quad + \frac{p_i}{N} \left(-\nu - \gamma \frac{p_i}{N} + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) \frac{d}{dp_i} f(z^N) + \frac{p_i}{N} \tilde{\beta}_P \frac{d^2}{dp_i^2} f(z^N) \\
 & \quad + \sum_{j \in \mathcal{D}} \tilde{\kappa}_H m(i, j) \left(\frac{a_j}{N} - \frac{a_i}{N} \right) \frac{d}{da_i} f(z^N) + \sum_{j \in \mathcal{D}} \tilde{\kappa}_H m(i, j) \left(\frac{c_j}{N} - \frac{c_i}{N} \right) \frac{d}{dc_i} f(z^N) \\
 & \quad \left. + \sum_{j \in \mathcal{D}} \tilde{\kappa}_P m(i, j) \left(\frac{p_j}{N} - \frac{p_i}{N} \right) \frac{d}{dp_i} f(z^N) + O\left(\frac{1}{N}\right) \right\}. \tag{4.6}
 \end{aligned}$$

For all $N \in \mathbb{N}$, all $f \in \text{Dom}(\mathcal{G})$, $z^N = \left(\frac{a_k}{N}, \frac{c_k}{N}, \frac{p_k}{N} \right)_{k \in \mathcal{D}} \in E_2^3$, such that $\mathbb{P}[Z_0^N = z^N] = 1$, it follows that

$$\begin{aligned} \mathcal{G}_N f(z^N) = & \sum_{i \in \mathcal{D}} \left\{ \left[\tilde{\kappa}_H \sum_{j \in \mathcal{D}} m(i, j) \left(\frac{a_j}{N} - \frac{a_i}{N} \right) + \frac{a_i}{N} \left(\lambda - \frac{\lambda}{K} \frac{a_i + c_i}{N} - \delta \frac{p_i}{N} \right. \right. \right. \\ & \left. \left. \left. - \tilde{\alpha} \right) \right] \frac{d}{da_i} f(z^N) + \frac{a_i}{N} \tilde{\beta}_H \frac{d^2}{da_i^2} f(z^N) + \left[\tilde{\kappa}_H \sum_{j \in \mathcal{D}} m(i, j) \left(\frac{c_j}{N} - \frac{c_i}{N} \right) + \frac{c_i}{N} \left(\lambda - \frac{\lambda}{K} \frac{a_i + c_i}{N} \right. \right. \right. \\ & \left. \left. \left. - \delta \frac{p_i}{N} \right) \right] \frac{d}{dc_i} f(z^N) + \frac{c_i}{N} \tilde{\beta}_H \frac{d^2}{dc_i^2} f(z^N) + \left[\tilde{\kappa}_P \sum_{j \in \mathcal{D}} m(i, j) \left(\frac{p_j}{N} - \frac{p_i}{N} \right) \right. \right. \\ & \left. \left. + \frac{p_i}{N} \left(-\nu - \gamma \frac{p_i}{N} + \eta \frac{c_i}{N} + (\eta - \rho) \frac{a_i}{N} \right) \right] \frac{d}{dp_i} f(z^N) + \frac{p_i}{N} \tilde{\beta}_P \frac{d^2}{dp_i^2} f(z^N) + O\left(\frac{1}{N}\right) \right\}. \end{aligned} \quad (4.7)$$

If for large N , it holds for all $i \in \mathcal{D}$ that $\frac{a_i}{N} \approx \tilde{A}_0(i) \in [0, \infty)$, $\frac{c_i}{N} \approx \tilde{C}_0(i) \in [0, \infty)$, and $\frac{p_i}{N} \approx \tilde{P}_0(i) \in [0, \infty)$, then the diffusion approximation of (4.1) takes the following form. Let $\tilde{W}^A(i), \tilde{W}^C(i), \tilde{W}^P(i): [0, \infty) \times \Omega \rightarrow \mathbb{R}$, $i \in \mathcal{D}$, be independent Brownian motions with continuous sample paths and let $\tilde{A}, \tilde{C}, \tilde{P}: [0, \infty) \times \mathcal{D} \times \Omega \rightarrow [0, \infty)$, be adapted processes with continuous sample paths that for all $i \in \mathcal{D}$ and all $t \in [0, \infty)$ satisfy \mathbb{P} -a.s.

$$\begin{aligned} \tilde{A}_t(i) = & \tilde{A}_0(i) + \int_0^t \left(\tilde{\kappa}_H \sum_{j \in \mathcal{D}} m(i, j) \left(\tilde{A}_s(j) - \tilde{A}_s(i) \right) + \tilde{A}_s(i) \left[\lambda \left(1 - \frac{\tilde{A}_s(i) + \tilde{C}_s(i)}{K} \right) \right. \right. \\ & \left. \left. - \delta \tilde{P}_s(i) - \tilde{\alpha} \right] \right) ds + \int_0^t \sqrt{\tilde{\beta}_H \tilde{A}_s(i)} d\tilde{W}_s^A(i), \\ \tilde{C}_t(i) = & \tilde{C}_0(i) + \int_0^t \left(\tilde{\kappa}_H \sum_{j \in \mathcal{D}} m(i, j) \left(\tilde{C}_s(j) - \tilde{C}_s(i) \right) + \tilde{C}_s(i) \left[\lambda \left(1 - \frac{\tilde{A}_s(i) + \tilde{C}_s(i)}{K} \right) \right. \right. \\ & \left. \left. - \delta \tilde{P}_s(i) \right] \right) ds + \int_0^t \sqrt{\tilde{\beta}_H \tilde{C}_s(i)} d\tilde{W}_s^C(i), \\ \tilde{P}_t(i) = & \tilde{P}_0(i) + \int_0^t \left(\tilde{\kappa}_P \sum_{j \in \mathcal{D}} m(i, j) \left(\tilde{P}_s(j) - \tilde{P}_s(i) \right) + \tilde{P}_s(i) \left[-\nu - \gamma \tilde{P}_s(i) \right. \right. \\ & \left. \left. + \eta \tilde{C}_s(i) + (\eta - \rho) \tilde{A}_s(i) \right] \right) ds + \int_0^t \sqrt{\tilde{\beta}_P \tilde{P}_s(i)} d\tilde{W}_s^P(i). \end{aligned} \quad (4.8)$$

Then, to obtain the classical assumptions of weak migration, weak selection, and weak genetic drift, the parameters are rescaled, i.e., $\tilde{\beta}_H$ is substituted by β_H^N , $\tilde{\beta}_P$ by β_P^N , $\tilde{\kappa}_H$ by κ_H^N , $\tilde{\kappa}_P$ by κ_P^N , and $\tilde{\alpha}$ by α^N , for $N \in \mathbb{N}$, such that $\lim_{N \rightarrow \infty} N\beta_H^N = \beta_H$, $\lim_{N \rightarrow \infty} N\beta_P^N = \beta_P$,

$\lim_{N \rightarrow \infty} N\kappa_H^N = \kappa_H$, $\lim_{N \rightarrow \infty} N\kappa_P^N = \kappa_P$, $\lim_{N \rightarrow \infty} N\alpha^N = \alpha$, for some $\beta_H, \beta_P, \kappa_H, \kappa_P$, and $\alpha \in (0, \infty)$. This results in the diffusion model that is analyzed in the remainder of Section 4.1.

4.1.3 Diffusion model

The following diffusion model forms the basis of the analysis in subsequent sections. It can be derived from the particle model in Section 2.1.1 via the diffusion approximation provided in Section 4.1.2. The derivations and results in the remainder of Section 4.1 are adapted from Hutzenthaler et al. (2015).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let \mathcal{D} be an at most countable set (the set of demes) and let $m \in [0, \infty)^{\mathcal{D} \times \mathcal{D}}$ satisfy for every $i \in \mathcal{D}$ that $\sum_{k \in \mathcal{D}} m(k, i) = \sum_{k \in \mathcal{D}} m(i, k) = 1$. Let $\lambda, K, \delta, \nu, \gamma, \eta, \rho \in (0, \infty)$ satisfy $\rho < \eta$. For every $N \in \mathbb{N}$, let $\kappa_H^N, \kappa_P^N, \alpha^N, \beta_H^N, \beta_P^N, \iota_H^N, \iota_P^N \in [0, \infty)$, let $W^{A,N}(i), W^{C,N}(i), W^{P,N}(i): [0, \infty) \times \Omega \rightarrow \mathbb{R}, i \in \mathcal{D}$, be independent Brownian motions with continuous sample paths, let $A^N, C^N, P^N: [0, \infty) \times \mathcal{D} \times \Omega \rightarrow [0, \infty)$ be adapted processes with continuous sample paths that for all $t \in [0, \infty)$ and all $i \in \mathcal{D}$ satisfy \mathbb{P} -a.s.

$$\begin{aligned}
 A_t^N(i) &= A_0^N(i) + \int_0^t \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (A_s^N(j) - A_s^N(i)) ds \\
 &\quad + \int_0^t A_s^N(i) \left[\lambda \left(1 - \frac{A_s^N(i) + C_s^N(i)}{K} \right) - \delta P_s^N(i) - \alpha^N \right] + \iota_H^N \frac{A_s^N(i)}{A_s^N(i) + C_s^N(i)} ds \\
 &\quad + \int_0^t \sqrt{\beta_H^N A_s^N(i)} dW_s^{A,N}(i), \\
 C_t^N(i) &= C_0^N(i) + \int_0^t \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (C_s^N(j) - C_s^N(i)) ds \\
 &\quad + \int_0^t C_s^N(i) \left[\lambda \left(1 - \frac{A_s^N(i) + C_s^N(i)}{K} \right) - \delta P_s^N(i) \right] + \iota_H^N \frac{C_s^N(i)}{A_s^N(i) + C_s^N(i)} ds \\
 &\quad + \int_0^t \sqrt{\beta_H^N C_s^N(i)} dW_s^{C,N}(i), \\
 P_t^N(i) &= P_0^N(i) + \int_0^t \kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) (P_s^N(j) - P_s^N(i)) ds \\
 &\quad + \int_0^t P_s^N(i) [-\nu - \gamma P_s^N(i) + \eta C_s^N(i) + (\eta - \rho) A_s^N(i)] + \iota_P^N ds \\
 &\quad + \int_0^t \sqrt{\beta_P^N P_s^N(i)} dW_s^{P,N}(i),
 \end{aligned} \tag{4.9}$$

let $H^N: [0, \infty) \times \mathcal{D} \times \Omega \rightarrow [0, \infty)$ satisfy $H^N = A^N + C^N$, and let $F^N: [0, \infty) \times \mathcal{D} \times \Omega \rightarrow$

$[0, 1]$ satisfy $F^N = \frac{A^N}{H^N}$. For every $N \in \mathbb{N}$, the process H^N describes the size of the host (or prey) population, P^N describes the size of the parasite (or predator) population, A^N describes the number of altruists, and C^N describes the number of cheaters, each measured in units of N individuals. Existence of solutions to (4.9), which is assumed here, can be established in suitable Liggett-Spitzer spaces if \mathcal{D} is an Abelian group and if m is translation invariant and irreducible, as shown in Proposition 2.1 in Hutzenthaler and Wakolbinger (2007).

Equation (4.9) extends the classical model for host-parasite interactions by Lotka (1920) and Volterra (1926) with stochastic fluctuations and it incorporates within-species competition and spatial structure. The immigration terms $\int_0^t \iota_H^N \frac{A_s^N(i)}{A_s^N(i) + C_s^N(i)} ds$, $\int_0^t \iota_H^N \frac{C_s^N(i)}{A_s^N(i) + C_s^N(i)} ds$, and $\int_0^t \iota_P^N ds$ are included for technical purposes, as they prevent the processes from going to extinction due to random fluctuations in the finite setting $N < \infty$, as shown in Lemma 4.1.3 and Lemma 4.1.4. In order to prove the results, some assumptions are necessary.

Assumption 4.1.1. *In the setting of Section 4.1.3 it holds that $\lambda > \nu$, $\eta - \rho > \frac{\lambda}{K}$, $\gamma \geq 2\delta$, for all $N \in \mathbb{N}$ it holds that $\alpha^N + \kappa_H^N \leq \frac{\lambda}{4}$, $\iota_P^N \leq \frac{\lambda(\nu + \lambda)}{8\delta}$, $\kappa_P^N + \kappa_H^N + \alpha^N \leq \frac{\lambda - \nu}{2}$, $\iota_H^N \geq \frac{4\delta\kappa_P^N}{3(\nu + \lambda)} + \frac{3}{2}\beta_H^N$, $\iota_P^N \geq \beta_P^N$, and there exist $\sigma = (\sigma_i)_{i \in \mathcal{D}} \in (0, \infty)^{\mathcal{D}}$ and $c \in (0, \infty)$ such that $\sum_{i \in \mathcal{D}} \sigma_i < \infty$, such that for every $j \in \mathcal{D}$ it holds that*

$$\sum_{i \in \mathcal{D}} \sigma_i m(i, j) \leq c\sigma_j, \quad (4.10)$$

and such that $\sup_{N \in \mathbb{N}} \mathbb{E} \left[\left\| \left(H_0^N + P_0^N \right)^4 + \frac{1}{(H_0^N)^2} + \frac{P_0^N}{(H_0^N)^2} + \frac{1}{P_0^N} + \frac{1}{P_0^N H_0^N} \right\|_{\sigma} \right] < \infty$.

4.1.4 Convergence of the relative frequency of altruists

In this section, the process describing the relative frequency of altruists among all hosts is derived by speeding up the processes with a factor of N in the limit $N \rightarrow \infty$.

4.1.4.1 Setting

Assume the setting of Section 4.1.3 and define $\bar{\kappa}_H := \sup_{N \in \mathbb{N}} \kappa_H^N$, $\bar{\kappa}_P := \sup_{N \in \mathbb{N}} \kappa_P^N$, $\bar{\beta}_H := \sup_{N \in \mathbb{N}} \beta_H^N$, $\bar{\beta}_P := \sup_{N \in \mathbb{N}} \beta_P^N$, $\bar{\iota}_H := \sup_{N \in \mathbb{N}} \iota_H^N$, $\bar{\iota}_P := \sup_{N \in \mathbb{N}} \iota_P^N$, and $\beta_H := \lim_{N \rightarrow \infty} \beta_H^N$. For all $z = (z_i)_{i \in \mathcal{D}} \in (0, \infty)^{\mathcal{D}}$ and $p \in \mathbb{R}$ let $z^p = (z_i^p)_{i \in \mathcal{D}}$. Furthermore, let $\underline{1} := (1)_{i \in \mathcal{D}} \in l_{\sigma}^1$. Define $E_1 := [0, 1]^{\mathcal{D}}$ and $E_2 := l_{\sigma}^1 \cap [0, \infty)^{\mathcal{D}}$. For all $i \in \mathcal{D}$ and all $N \in \mathbb{N}$ let $W^{H,N}(i): [0, \infty) \times \Omega \rightarrow \mathbb{R}$ and $W^{F,N}(i): [0, \infty) \times \Omega \rightarrow \mathbb{R}$ be stochastic processes with continuous sample paths such that for every $t \in [0, \infty)$ it holds \mathbb{P} -a.s. that

$$dW_t^{H,N}(i) = \frac{\sqrt{A_t^N(i)} dW_t^{A,N}(i) + \sqrt{C_t^N(i)} dW_t^{C,N}(i)}{\sqrt{H_t^N(i)}} \quad (4.11)$$

and

$$dW_t^{F,N}(i) = \frac{\sqrt{C_t^N(i)} dW_t^{A,N}(i) - \sqrt{A_t^N(i)} dW_t^{C,N}(i)}{\sqrt{H_t^N(i)}}, \quad (4.12)$$

respectively, with $W_0^{H,N}(i) = W_0^{F,N}(i) = 0$. Furthermore, define functions h_∞ and p_∞ by

$$\begin{aligned} [0, 1] \ni x \mapsto h_\infty(x) &:= \frac{K(\delta\nu + \gamma\lambda)}{\lambda\gamma + \delta K(\eta - \rho x)} = \frac{1}{b(a-x)} \in (0, \infty), \\ [0, 1] \ni x \mapsto p_\infty(x) &:= \frac{\lambda K(\eta - \rho x) - \lambda\nu}{\lambda\gamma + \delta K(\eta - \rho x)} = \frac{\lambda}{\delta} \left(1 - \frac{1}{Kb(a-x)} \right) \in (0, \infty), \end{aligned} \quad (4.13)$$

where $a := \frac{\lambda\gamma + \delta K\eta}{\delta K\rho}$ and $b := \frac{\delta\rho}{\delta\nu + \lambda\gamma}$. These functions are straightforward adaptations of the equilibrium values of standard Lotka-Volterra equations (cf. Theorem 3.1.1). Note that for the functions given by equation (4.13) to be well defined it is necessary that $Kb(a-1) > 1$ or, equivalently, that $K(\eta - \rho) > \nu$.

4.1.4.2 Preliminaries

Assume the setting of Section 4.1.4.1. In this section some first results are collected that are used in the proofs of the statements in subsequent sections.

Lemma 4.1.2. *Assume the setting of Section 4.1.4.1. Then $W^{H,N}(i)$ and $W^{F,N}(i)$, $N \in \mathbb{N}$, $i \in \mathcal{D}$, are independent Brownian motions and for all $t \in [0, \infty)$, all $i \in \mathcal{D}$, and all $N \in \mathbb{N}$ it \mathbb{P} -a.s. holds that*

$$H_t^N(i) = H_0^N(i) + \int_0^t \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_s^N(j) + (\lambda - \kappa_H^N - \alpha^N F_s^N(i)) H_s^N(i) \quad (4.14)$$

$$- \frac{\lambda}{K} (H_s^N(i))^2 - \delta P_s^N(i) H_s^N(i) + \iota_H^N ds + \int_0^t \sqrt{\beta_H^N H_s^N(i)} dW_s^{H,N}(i),$$

$$F_t^N(i) = F_0^N(i) + \int_0^t \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (F_s^N(j) - F_s^N(i)) \frac{H_s^N(j)}{H_s^N(i)} \quad (4.15)$$

$$- \alpha^N F_s^N(i) (1 - F_s^N(i)) ds + \int_0^t \sqrt{\frac{\beta_H^N F_s^N(i) (1 - F_s^N(i))}{H_s^N(i)}} dW_s^{F,N}(i),$$

$$P_t^N(i) = P_0^N(i) + \int_0^t \kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) P_s^N(j) - (\kappa_P^N + \nu) P_s^N(i) - \gamma (P_s^N(i))^2 \quad (4.16)$$

$$+ (\eta - \rho F_s^N(i)) P_s^N(i) H_s^N(i) + \iota_P^N ds + \int_0^t \sqrt{\beta_P^N P_s^N(i)} dW_s^{P,N}(i).$$

Proof. For all $t \in [0, \infty)$, all $N \in \mathbb{N}$, and all $i \in \mathcal{D}$ it holds that

$$\langle W^{H,N}(i), W^{F,N}(i) \rangle_t = \int_0^t \frac{\sqrt{A_s^N(i) C_s^N(i)} - \sqrt{A_s^N(i) C_s^N(i)}}{H_s^N(i)} ds = 0 \quad (4.17)$$

as well as $\langle W^{H,N}(i) \rangle_t = \langle W^{F,N}(i) \rangle_t = t$. Hence, $W^{H,N}(i)$ and $W^{F,N}(i)$, $N \in \mathbb{N}$, $i \in \mathcal{D}$, are independent Brownian motions. Equation (4.14) follows from Itô's lemma (e.g., Klenke, 2008) and rearranging terms. Furthermore, applying Itô's lemma it follows for all $t \in [0, \infty)$, all $i \in \mathcal{D}$, and all $N \in \mathbb{N}$ that \mathbb{P} -a.s.

$$\begin{aligned}
 F_t^N(i) &= F_0^N(i) + \int_0^t \frac{C_s^N(i)}{(H_s^N(i))^2} \left(\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (A_s^N(j) - A_s^N(i)) \right. \\
 &\quad \left. + A_s^N(i) \left(\lambda \left(1 - \frac{H_s^N(i)}{K} \right) - \delta P_s^N(i) - \alpha^N \right) + \iota_H^N \frac{A_s^N(i)}{H_s^N(i)} \right) ds \\
 &\quad + \int_0^t \frac{C_s^N(i)}{(H_s^N(i))^2} \sqrt{\beta_H^N A_s^N(i)} dW_s^A(i) - \int_0^t \frac{A_s^N(i)}{(H_s^N(i))^2} \left(\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (C_s^N(j) \right. \\
 &\quad \left. - C_s^N(i)) + C_s^N(i) \left(\lambda \left(1 - \frac{H_s^N(i)}{K} \right) - \delta P_s^N(i) \right) + \iota_H^N \frac{C_s^N(i)}{H_s^N(i)} \right) ds \\
 &\quad - \int_0^t \frac{A_s^N(i)}{(H_s^N(i))^2} \sqrt{\beta_H^N C_s^N(i)} dW_s^C(i) - \int_0^t \frac{C_s^N(i)}{(H_s^N(i))^3} \beta_H^N A_s^N(i) + \frac{A_s^N(i)}{(H_s^N(i))^3} \beta_H^N C_s^N(i) ds \\
 &= F_0^N(i) + \int_0^t \frac{\kappa_H^N}{H_s^N(i)} \sum_{j \in \mathcal{D}} m(i, j) \left((1 - F_s^N(i)) F_s^N(j) H_s^N(j) \right. \\
 &\quad \left. - F_s^N(i) (1 - F_s^N(j)) H_s^N(j) \right) - \alpha^N F_s^N(i) (1 - F_s^N(i)) ds \\
 &\quad + \int_0^t \sqrt{\frac{\beta_H^N F_s^N(i) (1 - F_s^N(i))}{H_s^N(i)}} dW_s^{F,N}(i)
 \end{aligned} \tag{4.18}$$

and (4.15) follows. Finally, the definitions of $(H^N)_{N \in \mathbb{N}}$ and $(F^N)_{N \in \mathbb{N}}$ yield (4.16). \square

Lemma 4.1.3. *Assume the setting of Section 4.1.4.1 and assume that for all $N \in \mathbb{N}$ it holds that $\iota_H^N \geq \frac{1}{2} \beta_H^N$. Furthermore, assume for all $N \in \mathbb{N}$ and all $i \in \mathcal{D}$ that \mathbb{P} -a.s. $H_0^N(i) > 0$. Then it holds that*

$$\mathbb{P} [H_u^N(i) > 0, \text{ for all } u \in [0, \infty), \text{ all } N \in \mathbb{N}, \text{ and all } i \in \mathcal{D}] = 1. \tag{4.19}$$

Proof. For every $N, M \in \mathbb{N}$ let $\hat{H}^{N,M}: [0, \infty) \times \mathcal{D} \times \Omega \rightarrow [0, \infty)$ be an adapted process with continuous sample paths that for all $t \in [0, \infty)$ and all $i \in \mathcal{D}$ satisfies \mathbb{P} -a.s.

$$\begin{aligned}
 \hat{H}_t^{N,M}(i) &= \hat{H}_0^{N,M}(i) + \int_0^t \left[\hat{H}_s^{N,M}(i) \left(\lambda - \alpha^N - \kappa_H^N - \frac{\lambda}{K} \hat{H}_s^{N,M}(i) - \delta M \right) \right. \\
 &\quad \left. + \iota_H^N \right] ds + \int_0^t \sqrt{\beta_H^N \hat{H}_s^{N,M}(i)} dW_s^{H,N}(i)
 \end{aligned} \tag{4.20}$$

with $\hat{H}_0^{N,M}(i) = H_0^N(i)$. Due to Feller's boundary classification (e.g., Ethier and Kurtz, 1986, p. 366) with the assumption that for all $N \in \mathbb{N}$ it holds that $\iota_H^N \geq \frac{1}{2}\beta_H^N$ it follows for every $N, M \in \mathbb{N}$ and all $i \in \mathcal{D}$ that

$$\mathbb{P} \left[\hat{H}_t^{N,M}(i) > 0, \text{ for all } t \in [0, \infty) \right] = 1. \quad (4.21)$$

For all $N, M \in \mathbb{N}$, all $i \in \mathcal{D}$, and all $t \in [0, \infty)$ consider the event $A_M^N(i) := \left\{ \sup_{s \in [0, t]} P_s^N(i) \leq M \right\}$. It holds for all $N, M \in \mathbb{N}$, all $i \in \mathcal{D}$, and all $t \in [0, \infty)$ that

$$\begin{aligned} A_M^N(i) &\subseteq A_{M+1}^N(i), \\ \mathbb{P} \left[\bigcup_{M \in \mathbb{N}} A_M^N(i) \right] &= \mathbb{P} \left[\sup_{s \in [0, t]} P_s^N(i) < \infty \right] = 1. \end{aligned} \quad (4.22)$$

Using a comparison result due to Ikeda and Watanabe (see e.g., Rogers and Williams, 2000, Theorem V.43.1), implies for all $N, M \in \mathbb{N}$, all $i \in \mathcal{D}$, and all $t \in [0, \infty)$ that

$$\mathbb{P} \left[\exists u \in [0, t] : H_u^N(i) < \hat{H}_u^{N,M}(i), \sup_{s \in [0, t]} P_s^N(i) \leq M \right] = 0. \quad (4.23)$$

Thus, combining (4.21), (4.22), and (4.23) it holds for all $N \in \mathbb{N}$, all $i \in \mathcal{D}$, and all $t \in [0, \infty)$ that

$$\begin{aligned} 1 &\geq \mathbb{P} \left[H_u^N(i) > 0, \text{ for all } u \in [0, t] \right] = 1 - \mathbb{P} \left[\exists u \in [0, t] : H_u^N(i) = 0 \right] \\ &\geq 1 - \sum_{M \in \mathbb{N}} \mathbb{P} \left[\exists u \in [0, t] : H_u^N(i) = 0, \sup_{s \in [0, t]} P_s^N(i) \leq M \right] \\ &\geq 1 - \sum_{M \in \mathbb{N}} \mathbb{P} \left[\exists u \in [0, t] : H_u^N(i) < \hat{H}_u^{N,M}(i), \sup_{s \in [0, t]} P_s^N(i) \leq M \right] = 1. \end{aligned} \quad (4.24)$$

This implies for all $N \in \mathbb{N}$, all $i \in \mathcal{D}$, and all $t \in [0, \infty)$ that $\mathbb{P} \left[H_u^N(i) > 0, \text{ for all } u \in [0, t] \right] = 1$, which in turn implies (4.19). This completes the proof of Lemma 4.1.3. \square

Lemma 4.1.4. *Assume the setting of Section 4.1.4.1 and assume that for all $N \in \mathbb{N}$ it holds that $\iota_P^N \geq \frac{1}{2}\beta_P^N$. Furthermore, assume for all $N \in \mathbb{N}$ and all $i \in \mathcal{D}$ that \mathbb{P} -a.s. $P_0^N(i) > 0$. Then it holds that*

$$\mathbb{P} \left[P_t^N(i) > 0, \text{ for all } t \in [0, \infty), \text{ all } N \in \mathbb{N}, \text{ and all } i \in \mathcal{D} \right] = 1. \quad (4.25)$$

Proof. Analogous to the proof of Lemma 4.1.3. \square

Lemma 4.1.5. *Assume the setting of Section 4.1.4.1. For all $x = (x_i)_{i \in \mathcal{D}} \in E_2$, all $p \in [1, \infty)$, and all sets $\mathcal{D}' \subseteq \mathcal{D}$ it holds that*

$$\sum_{i \in \mathcal{D}'} \sigma_i \left(\sum_{j \in \mathcal{D}} m(i, j) x_j \right)^p \leq \sum_{i \in \mathcal{D}} c \sigma_i x_i^p. \quad (4.26)$$

Proof. For any $x = (x_i)_{i \in \mathcal{D}} \in E_2$, any $p \in [1, \infty)$, and any set $\mathcal{D}' \subseteq \mathcal{D}$ it follows from Jensen's inequality and (4.10) that

$$\sum_{i \in \mathcal{D}'} \sigma_i \left(\sum_{j \in \mathcal{D}} m(i, j) x_j \right)^p \leq \sum_{i \in \mathcal{D}} \sigma_i \sum_{j \in \mathcal{D}} m(i, j) x_j^p \leq \sum_{i \in \mathcal{D}} c \sigma_i x_i^p. \quad (4.27)$$

\square

4.1.4.3 Strong convergence of the spatial stochastic Lotka-Volterra processes

In this section the convergence of the time-rescaled Lotka-Volterra processes as given in (4.14) and (4.16) will be shown. In Lemmas 4.1.6, 4.1.7, and 4.1.8 bounds for the expected value of the sum (over sets of demes) of functionals of the processes weighted by σ will be provided. These are then used in Theorem 4.1.9 to show a result on the behavior of a spatial analogue of a well-known Lyapunov function (e.g., Dobrinevski and Frey, 2012). From that the convergence of the processes follows immediately in Theorem 3.1.1.

Lemma 4.1.6. *Assume the setting of Section 4.1.4.1, let $p \in \{1\} \cup [2, \infty)$, and define $\bar{c} := \frac{4 \min \left\{ \frac{1}{2\eta} \frac{\lambda}{K}, \frac{1}{4}, \frac{1}{\delta} \gamma \right\} \left[2\eta \bar{\ell}_H + \delta \bar{\ell}_P + (p-1) \left(2\eta \bar{\beta}_H + \frac{1}{2} \delta \bar{\beta}_P \right) \right]}{\left(\lambda + \left(1 - \frac{1}{p} + \frac{c}{p} \right) (\bar{\kappa}_H + \bar{\kappa}_P) \right)^2}$. Then it holds that*

$$\begin{aligned} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\left\| (2\eta H_t^N + \delta P_t^N)^p \right\|_\sigma \right] &\leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\left\| (2\eta H_0^N + \delta P_0^N)^p \right\|_\sigma \right] \\ &+ \left\| \mathbb{1} \right\|_\sigma \left(\frac{\lambda + \left(1 - \frac{1}{p} + \frac{c}{p} \right) (\bar{\kappa}_H + \bar{\kappa}_P)}{2 \min \left\{ \frac{1}{2\eta} \frac{\lambda}{K}, \frac{1}{4}, \frac{1}{\delta} \gamma \right\}} \right)^p \left(1 + \sqrt{1 + \bar{c}} \right)^p. \end{aligned} \quad (4.28)$$

Proof. Under the assumption that $\sup_{N \in \mathbb{N}} \mathbb{E} \left[\left\| (H_0^N + P_0^N)^p \right\|_\sigma \right] = \infty$, the claim trivially holds. For the remainder of the proof, assume $\sup_{N \in \mathbb{N}} \mathbb{E} \left[\left\| (H_0^N + P_0^N)^p \right\|_\sigma \right] < \infty$. Define $\mathcal{D}_0 := \emptyset$ and for every $n \in \mathbb{N}$ let $\mathcal{D}_n \subseteq \mathcal{D}$ be a set with $|\mathcal{D}_n| = \min \{n, |\mathcal{D}|\}$ and $\mathcal{D}_n \supseteq \mathcal{D}_{n-1}$. Define real numbers $c_0 := \min \left\{ \frac{1}{2\eta} \frac{\lambda}{K}, \frac{1}{4}, \frac{1}{\delta} \gamma \right\}$, $c_1 := p \left[2\eta \bar{\ell}_H + \delta \bar{\ell}_P + (p-1) \left(2\eta \bar{\beta}_H + \frac{1}{2} \delta \bar{\beta}_P \right) \right] \in (0, \infty)$, $c_2 := \lambda p + (p-1+c)(\bar{\kappa}_H + \bar{\kappa}_P) \in (0, \infty)$, $c_3 := c_0 p \left(\sum_{k \in \mathcal{D}} \sigma_k \right)^{-\frac{1}{p}} \in (0, \infty)$, and

$c_4 := c_1 \left(\sum_{k \in \mathcal{D}} \sigma_k \right)^{\frac{1}{p}} \in (0, \infty)$. For all $N \in \mathbb{N}$, $t \in [0, \infty)$ define $Y_t^N := 2\eta H_t^N + \delta P_t^N$ and for all $N, n \in \mathbb{N}$ and all $t \in [0, \infty)$ let $M_t^{N,n}$ be a real-valued random variable such that \mathbb{P} -a.s. it holds that

$$M_t^{N,n} = \sum_{i \in \mathcal{D}_n} \sigma_i \left(\int_0^t 2\eta p (Y_u^N(i))^{p-1} \sqrt{\beta_H^N H_u^N(i)} dW_u^{H,N}(i) + \int_0^t \delta p (Y_u^N(i))^{p-1} \sqrt{\beta_P^N P_u^N(i)} dW_u^{P,N}(i) \right). \quad (4.29)$$

Applying Itô's lemma it follows for all $N, n \in \mathbb{N}$ and all $t \in [0, \infty)$ that \mathbb{P} -a.s.

$$\begin{aligned} & \sum_{i \in \mathcal{D}_n} \sigma_i (Y_t^N(i))^p - \sum_{i \in \mathcal{D}_n} \sigma_i (Y_0^N(i))^p \\ &= \sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t 2\eta p (Y_u^N(i))^{p-1} \left(\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) \right. \\ & \quad \left. + (\lambda - \kappa_H^N - \alpha^N F_u^N(i)) H_u^N(i) - \frac{\lambda}{K} (H_u^N(i))^2 - \delta P_u^N(i) H_u^N(i) + \iota_H^N \right) \\ & \quad + \delta p (Y_u^N(i))^{p-1} \left(\kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) - (\kappa_P^N + \nu) P_u^N(i) - \gamma (P_u^N(i))^2 \right. \\ & \quad \left. + (\eta - \rho F_u^N(i)) P_u^N(i) H_u^N(i) + \iota_P^N \right) + \frac{1}{2} 4\eta^2 p(p-1) (Y_u^N(i))^{p-2} \beta_H^N H_u^N(i) \\ & \quad + \frac{1}{2} \delta^2 p(p-1) (Y_u^N(i))^{p-2} \beta_P^N P_u^N(i) du + M_t^{N,n}. \end{aligned} \quad (4.30)$$

Because $1 \geq c_0 4$, $\frac{\lambda}{K} \geq c_0 2\eta$, and $\gamma \geq \delta c_0$ it holds for all $N, n \in \mathbb{N}$ and all $t \in [0, \infty)$ that

\mathbb{P} -a.s.

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}_n} \sigma_i (Y_t^N(i))^p - \sum_{i \in \mathcal{D}_n} \sigma_i (Y_0^N(i))^p \\
 & \leq \sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t p (Y_u^N(i))^{p-1} \left(2\eta \bar{\kappa}_H \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + 2\eta \lambda H_u^N(i) \right. \\
 & \quad \left. - c_0 (2\eta H_u^N(i))^2 - [\eta \delta + c_0 4\eta \delta] P_u^N(i) H_u^N(i) + 2\eta \bar{\iota}_H \right) \\
 & \quad + p (Y_u^N(i))^{p-1} \left(\delta \bar{\kappa}_P \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) + \lambda \delta P_u^N(i) - c_0 (\delta P_u^N(i))^2 \right. \\
 & \quad \left. + \eta \delta P_u^N(i) H_u^N(i) + \delta \bar{\iota}_P \right) + p(p-1) (Y_u^N(i))^{p-2} \left((2\eta \bar{\beta}_H \right. \\
 & \quad \left. + \frac{1}{2} \delta \bar{\beta}_P) 2\eta H_u^N(i) + (\frac{1}{2} \delta \bar{\beta}_P + 2\eta \bar{\beta}_H) \delta P_u^N(i) \right) du + M_t^{N,n} \\
 & = \int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i p (Y_u^N(i))^{p-1} 2\eta \bar{\kappa}_H \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + \sum_{i \in \mathcal{D}_n} \sigma_i p (Y_u^N(i))^{p-1} \\
 & \quad \left(\lambda (Y_u^N(i)) - c_0 (Y_u^N(i))^2 + (2\eta \bar{\iota}_H + \delta \bar{\iota}_P) \right) \\
 & \quad + \sum_{i \in \mathcal{D}_n} \sigma_i p (Y_u^N(i))^{p-1} \delta \bar{\kappa}_P \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) \\
 & \quad + \sum_{i \in \mathcal{D}_n} \sigma_i p(p-1) (2\eta \bar{\beta}_H + \frac{1}{2} \delta \bar{\beta}_P) (Y_u^N(i))^{p-1} du + M_t^{N,n}.
 \end{aligned} \tag{4.31}$$

Using Young's inequality and Lemma 4.1.5 it holds for all $N, n \in \mathbb{N}$ and all $t \in [0, \infty)$ that

\mathbb{P} -a.s.

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}_n} \sigma_i (Y_t^N(i))^p - \sum_{i \in \mathcal{D}_n} \sigma_i (Y_0^N(i))^p \\
 & \leq \int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i \frac{p-1}{p} p (Y_u^N(i))^p \bar{\kappa}_H + \sum_{i \in \mathcal{D}} \sigma_i \frac{1}{p} p \bar{\kappa}_H c (2\eta H_u^N(i))^p \\
 & \quad + \sum_{i \in \mathcal{D}_n} \sigma_i \lambda p (Y_u^N(i))^p + \sum_{i \in \mathcal{D}_n} \sigma_i c_1 (Y_u^N(i))^{p-1} - \sum_{i \in \mathcal{D}_n} \sigma_i c_0 p (Y_u^N(i))^{p+1} \\
 & \quad + \sum_{i \in \mathcal{D}_n} \sigma_i \frac{p-1}{p} p (Y_u^N(i))^p \bar{\kappa}_P + \sum_{i \in \mathcal{D}} \sigma_i \frac{1}{p} p \bar{\kappa}_P c (\delta P_u^N(i))^p du + M_t^{N,n} \\
 & \leq \int_0^t \sum_{i \in \mathcal{D}} \sigma_i c_2 (Y_u^N(i))^p + \sum_{i \in \mathcal{D}_n} \sigma_i c_1 (Y_u^N(i))^{p-1} - \sum_{i \in \mathcal{D}_n} \sigma_i c_0 p (Y_u^N(i))^{p+1} du \\
 & \quad + M_t^{N,n}.
 \end{aligned} \tag{4.32}$$

For $N, n, l \in \mathbb{N}$ define $[0, \infty]$ -valued stopping times

$$\tau_l^{N,n} := \inf \left(\left\{ t \in [0, \infty) : \sum_{i \in \mathcal{D}_n} \sigma_i (Y_t^N(i))^p > l \right\} \cup \infty \right). \tag{4.33}$$

Now for all $N, n, l \in \mathbb{N}$ and all $t \in [0, \infty)$ it holds that

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i \left[\left(2\eta p (Y_u^N(i))^{p-1} \sqrt{\beta_H^N H_u^N(i)} \right)^2 \right. \right. \\
 & \quad \left. \left. + \left(\delta p (Y_u^N(i))^{p-1} \sqrt{\beta_P^N P_u^N(i)} \right)^2 \right] du \right] \\
 & = \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i p^2 \left[2\eta \beta_H^N \left((Y_u^N(i))^{p-1} \sqrt{2\eta H_u^N(i)} \right)^2 \right. \right. \\
 & \quad \left. \left. + \delta \beta_P^N \left((Y_u^N(i))^{p-1} \sqrt{\delta P_u^N(i)} \right)^2 \right] du \right] \\
 & \leq (2\eta \beta_H^N + \delta \beta_P^N) \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i p^2 \left[\left(2 (Y_u^N(i))^{\frac{2p-1}{2}} \right)^2 \right] du \right]
 \end{aligned} \tag{4.34}$$

Young's inequality implies for all $N, n, l \in \mathbb{N}$ and all $t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i \left[\left(2\eta p (Y_u^N(i))^{p-1} \sqrt{\beta_H^N H_u^N(i)} \right)^2 \right. \right. \\
 & \quad \left. \left. + \left(\delta p (Y_u^N(i))^{p-1} \sqrt{\beta_P^N P_u^N(i)} \right)^2 \right] du \right] \\
 & \leq (2\eta\beta_H^N + \delta\beta_P^N) \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i \left[4p^2 \left(\frac{2p-1}{2p} (Y_u^N(i))^p + \frac{1}{2p} \right)^2 \right] du \right] \\
 & \leq (2\eta\beta_H^N + \delta\beta_P^N) \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \frac{\sigma_i^2}{\min_{k \in \mathcal{D}_n} \sigma_k} \left[((2p-1) (Y_u^N(i))^p + 1)^2 \right] du \right] \quad (4.35) \\
 & \leq \frac{2\eta\beta_H^N + \delta\beta_P^N}{\min_{k \in \mathcal{D}_n} \sigma_k} \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \left[\left(\sum_{i \in \mathcal{D}_n} \sigma_i ((2p-1) (Y_u^N(i))^p + 1) \right)^2 \right] du \right] \\
 & \leq \frac{2\eta\beta_H^N + \delta\beta_P^N}{\min_{k \in \mathcal{D}_n} \sigma_k} \mathbb{E} \left[\int_0^t \left[\left((2p-1) \sum_{i \in \mathcal{D}_n} \sigma_i (Y_{u \wedge \tau_l^{N,n}}^N(i))^p + \|\mathbf{1}\|_\sigma \right)^2 \right] du \right] \\
 & \leq \frac{2\eta\beta_H^N + \delta\beta_P^N}{\min_{k \in \mathcal{D}_n} \sigma_k} t \left[((2p-1)l + \|\mathbf{1}\|_\sigma)^2 \right] < \infty.
 \end{aligned}$$

Hence, for all $N, n, l \in \mathbb{N}$ and all $t \in [0, \infty)$ it holds that $\mathbb{E} \left[M_{t \wedge \tau_l^{N,n}}^{N,n} \right] = 0$. This together with (4.32) and Tonelli's theorem implies for all $N, n, l \in \mathbb{N}$ and all $t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (Y_{t \wedge \tau_l^{N,n}}^N(i))^p + \int_0^{t \wedge \tau_l^{N,n}} c_0 p \sum_{i \in \mathcal{D}_n} \sigma_i (Y_u^N(i))^{p+1} du \right] \\
 & \leq \mathbb{E} [\| (Y_0^N)^p \|_\sigma] + \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} c_2 \| (Y_u^N)^p \|_\sigma + c_1 \| (Y_u^N)^{p-1} \|_\sigma du \right] \quad (4.36) \\
 & \leq \mathbb{E} [\| (Y_0^N)^p \|_\sigma] + \int_0^t c_2 \mathbb{E} [\| (Y_u^N)^p \|_\sigma] + c_1 \mathbb{E} [\| (Y_u^N)^{p-1} \|_\sigma] du.
 \end{aligned}$$

For every $N, n \in \mathbb{N}$ the map $[0, \infty) \ni t \mapsto \sum_{i \in \mathcal{D}_n} \sigma_i (Y_t^N(i))^p \in \mathbb{R}$ is \mathbb{P} -a.s. continuous which implies for all $N, n \in \mathbb{N}$ and all $t \in [0, \infty)$ that $\mathbb{P} \left[\lim_{l \rightarrow \infty} \tau_l^{N,n} < t \right] = 0$. From Tonelli's theorem and monotone convergence, then using Fatou's lemma, and finally apply-

ing (4.36) it follows for all $N \in \mathbb{N}$ and all $t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_t^N(i))^p \right] + \int_0^t c_0 p \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^{p+1} \right] du \\
 &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (Y_t^N(i))^p + \int_0^t c_0 p \sum_{i \in \mathcal{D}_n} \sigma_i (Y_u^N(i))^{p+1} du \right] \\
 &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\lim_{l \rightarrow \infty} \left(\sum_{i \in \mathcal{D}_n} \sigma_i (Y_{t \wedge \tau_l^{N,n}}^N(i))^p + \int_0^{t \wedge \tau_l^{N,n}} c_0 p \sum_{i \in \mathcal{D}_n} \sigma_i (Y_u^N(i))^{p+1} du \right) \right] \quad (4.37) \\
 &\leq \lim_{n \rightarrow \infty} \liminf_{l \rightarrow \infty} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (Y_{t \wedge \tau_l^{N,n}}^N(i))^p + \int_0^{t \wedge \tau_l^{N,n}} c_0 p \sum_{i \in \mathcal{D}_n} \sigma_i (Y_u^N(i))^{p+1} du \right] \\
 &\leq \mathbb{E} [\| (Y_0^N)^p \|_\sigma] + \int_0^t c_2 \mathbb{E} [\| (Y_u^N)^p \|_\sigma] + c_1 \mathbb{E} [\| (Y_u^N)^{p-1} \|_\sigma] du.
 \end{aligned}$$

This implies using Jensen's inequality for all $N \in \mathbb{N}$ and all $t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} [\| (Y_t^N)^p \|_\sigma] - \mathbb{E} [\| (Y_0^N)^p \|_\sigma] \\
 &\leq \int_0^t c_2 \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^p \right] + c_1 \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^{p-1} \right] \\
 &\quad - c_0 p \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^{p+1} \right] du \\
 &= \int_0^t c_2 \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^p \right] + \frac{\sum_{k \in \mathcal{D}} \sigma_k}{\sum_{l \in \mathcal{D}} \sigma_l} \left(c_1 \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^{p-1} \right] \right. \\
 &\quad \left. - c_0 p \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^{p+1} \right] \right) du \quad (4.38) \\
 &\leq \int_0^t c_2 \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^p \right] + c_4 \left(\mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^p \right] \right)^{\frac{p-1}{p}} \\
 &\quad - c_3 \left(\mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (Y_u^N(i))^p \right] \right)^{\frac{p+1}{p}} du \\
 &= \int_0^t (\mathbb{E} [\| (Y_u^N)^p \|_\sigma])^{\frac{p-1}{p}} \left\{ c_4 + c_2 (\mathbb{E} [\| (Y_u^N)^p \|_\sigma])^{\frac{1}{p}} \right. \\
 &\quad \left. - c_3 (\mathbb{E} [\| (Y_u^N)^p \|_\sigma])^{\frac{2}{p}} \right\} du.
 \end{aligned}$$

For every $N \in \mathbb{N}$ let $z^N: [0, \infty) \rightarrow \mathbb{R}$ be a process that for all $t \in [0, \infty)$ satisfies

$$z_t^N = z_0^N + \int_0^t (z_s^N)^{\frac{p-1}{p}} \left\{ c_4 + c_2 (z_s^N)^{\frac{1}{p}} - c_3 (z_s^N)^{\frac{2}{p}} \right\} ds \quad (4.39)$$

with $z_0^N = \mathbb{E} [\| (Y_0^N)^p \|_\sigma]$, where uniqueness follows from local Lipschitz continuity. Using classical comparison results from the theory of ODEs, the above computation shows for all $N \in \mathbb{N}$ and all $t \in [0, \infty)$ that $\mathbb{E} [\| (Y_t^N)^p \|_\sigma] \leq z_t^N$ and for all $N \in \mathbb{N}$ that $\sup_{t \in [0, \infty)} z_t^N = \max \left\{ \mathbb{E} [\| (Y_0^N)^p \|_\sigma], \left(\frac{c_2}{2c_3} + \sqrt{\frac{(c_2)^2}{4c_3^2} + \frac{c_4}{c_3}} \right)^p \right\}$. Hence, it holds that

$$\begin{aligned} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} [\| (2\eta H_t^N + \delta P_t^N)^p \|_\sigma] &\leq \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} z_t^N \leq \sup_{N \in \mathbb{N}} \mathbb{E} [\| (Y_0^N)^p \|_\sigma] \\ &+ \left(\frac{c_2}{2c_3} + \sqrt{\frac{c_2^2}{4c_3^2} + \frac{c_4}{c_3}} \right)^p = \sup_{N \in \mathbb{N}} \mathbb{E} [\| (Y_0^N)^p \|_\sigma] + \frac{c_2}{2c_3} \left(1 + \sqrt{1 + \frac{c_4 c_3}{c_2^2}} \right)^p \\ &= \sup_{N \in \mathbb{N}} \mathbb{E} [\| (2\eta H_0^N + \delta P_0^N)^p \|_\sigma] + \|\mathbf{1}\|_\sigma \left(\frac{\lambda + (1 - \frac{1}{p} + \frac{c}{p})(\bar{\kappa}_H + \bar{\kappa}_P)}{2 \min \left\{ \frac{1}{2\eta}, \frac{\lambda}{K}, \frac{1}{4}, \frac{1}{\delta} \gamma \right\}} \right)^p \left(1 \right. \\ &\quad \left. + \sqrt{1 + \frac{4 \min \left\{ \frac{1}{2\eta}, \frac{\lambda}{K}, \frac{1}{4}, \frac{1}{\delta} \gamma \right\} [2\eta \bar{\iota}_H + \delta \bar{\iota}_P + (p-1)(2\eta \bar{\beta}_H + \frac{1}{2} \delta \bar{\beta}_P)]}{(\lambda + (1 - \frac{1}{p} + \frac{c}{p})(\bar{\kappa}_H + \bar{\kappa}_P))^2}} \right)^p. \end{aligned} \quad (4.40)$$

This completes the proof of Lemma 4.1.6. \square

Lemma 4.1.7. *Assume the setting of Section 4.1.4.1 and assume $\gamma \geq 2\delta$. Furthermore, assume for all $N \in \mathbb{N}$ that $\alpha^N + \kappa_H^N \leq \frac{\lambda}{4}$, $\iota_P^N \leq \frac{\lambda(\nu + \lambda)}{8\delta}$, and $\iota_H^N \geq \frac{4\delta\kappa_P^N}{3(\nu + \lambda)} + \frac{3}{2}\beta_H^N$. Let $\hat{\mathcal{D}} \subseteq \mathcal{D}$ be a set. Then it holds that*

$$\begin{aligned} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(\frac{2}{\lambda + \nu} \frac{P_t^N(i)}{(H_t^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_t^N(i))^2} \right) \right] &\leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(\frac{2}{\lambda + \nu} \frac{P_0^N}{(H_0^N)^2} \right. \right. \\ &\quad \left. \left. + \frac{1}{2\delta} \frac{1}{(H_0^N)^2} \right) \right] + \frac{4\bar{\kappa}_P c}{3\lambda(\lambda + \nu)} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i (P_t^N(i))^3 \right] \\ &+ \frac{4}{\lambda(\lambda + \nu)} \left(\frac{\eta^2}{\lambda} + \frac{4\lambda}{K^2} \right) \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i P_t^N(i) \right] + \frac{2}{K^2 \delta}. \end{aligned} \quad (4.41)$$

Proof. If the right-hand side of (4.41) is infinite, then the claim trivially holds. For the remainder of the proof assume the right-hand side of (4.41) to be finite. Define $\mathcal{D}_0 := \emptyset$ and for every $n \in \mathbb{N}$ let $\mathcal{D}_n \subseteq \hat{\mathcal{D}}$ be a set with $|\mathcal{D}_n| = \min \{n, |\hat{\mathcal{D}}|\}$ and $\mathcal{D}_n \supseteq \mathcal{D}_{n-1}$. Define

$c_1 := \frac{1}{\lambda + \nu}$ and for all $n \in \mathbb{N}$ let $c_0^n := \frac{2c_1 \bar{\kappa}_{PC}}{3} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (P_t^N(i))^3 \right] + 2c_1 \left[\frac{\eta^2}{\lambda} + \frac{4\lambda}{K^2} \right] \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i P_t^N(i) \right] + \frac{\lambda}{K^2 \delta}$. For $N, n, l \in \mathbb{N}$ define $[0, \infty]$ -valued stopping times

$$\tau_l^{N,n} := \inf \left(\left\{ t \in [0, \infty) : \sum_{i \in \mathcal{D}_n} \sigma_i \left(P_t^N(i) + (H_t^N(i))^{-1} \right) > l \right\} \cup \infty \right). \quad (4.42)$$

Lemma 4.1.3 yields that for all $N, n \in \mathbb{N}$ the map $[0, \infty) \ni t \mapsto \sum_{i \in \mathcal{D}_n} \sigma_i \left(P_t^N(i) + (H_t^N(i))^{-1} \right) \in \mathbb{R}$ is \mathbb{P} -a.s. continuous. Thereby, for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ it holds that

$$\mathbb{P} \left[\lim_{l \rightarrow \infty} \tau_l^{N,n} < t \right] = 0. \quad (4.43)$$

For all $t \in [0, \infty)$, $N, n, l \in \mathbb{N}$, Young's inequality implies

$$\begin{aligned} & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left(2c_1 \frac{\sqrt{\beta_P^N P_u^N(i)}}{(H_u^N(i))^2} \right)^2 du \right] \\ & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \frac{\sigma_i^5}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} t \sup_{u \in [0, t]} 4c_1^2 \bar{\beta}_P \left(\frac{1}{5} \left(P_{u \wedge \tau_l^{N,n}}^N(i) \right)^5 + \frac{4}{5} \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-5} \right) \right] \\ & \leq \frac{t 4c_1^2 \bar{\beta}_P}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} \mathbb{E} \left[\sup_{u \in [0, t]} \left(\sum_{i \in \mathcal{D}_n} \sigma_i \left(P_{u \wedge \tau_l^{N,n}}^N(i) + \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} \right)^5 \right) \right] \\ & \leq \frac{t 4c_1^2 \bar{\beta}_P}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} l^5 < \infty \end{aligned} \quad (4.44)$$

and

$$\begin{aligned} & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left(\left(4c_1 P_u^N(i) + \frac{1}{\delta} \right) \frac{\sqrt{\beta_H^N H_u^N(i)}}{(H_u^N(i))^3} \right)^2 du \right] \\ & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \frac{\sigma_i^7 t \sup_{u \in [0, t]} (4c_1 + \frac{1}{\delta})^2 \bar{\beta}_H}{\min_{k \in \mathcal{D}_n} \{\sigma_k^6\}} \left(\frac{2}{7} \left(P_{u \wedge \tau_l^{N,n}}^N(i) + 1 \right)^7 + \frac{5}{7} \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-7} \right) \right] \\ & \leq \frac{t (4c_1 + \frac{1}{\delta})^2 \bar{\beta}_H}{\min_{k \in \mathcal{D}_n} \{\sigma_k^6\}} \mathbb{E} \left[\sup_{u \in [0, t]} \left(\sum_{i \in \mathcal{D}_n} \sigma_i \left(P_{u \wedge \tau_l^{N,n}}^N(i) + 1 + \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} \right)^7 \right) \right] \\ & \leq \frac{t (4c_1 + \frac{1}{\delta})^2 \bar{\beta}_H}{\min_{k \in \mathcal{D}_n} \{\sigma_k^6\}} (l + \|\mathbf{1}\|_\sigma)^7 < \infty. \end{aligned} \quad (4.45)$$

Hence, for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ it follows that

$$\begin{aligned} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} 2c_1 \frac{\sqrt{\beta_P^N P_u^N(i)}}{(H_u^N(i))^2} dW_u^{P,N}(i) \right] &= 0, \\ \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left(4c_1 P_u^N(i) + \frac{1}{\delta} \right) \frac{\sqrt{\beta_H^N H_u^N(i)}}{(H_u^N(i))^3} dW_u^{H,N}(i) \right] &= 0. \end{aligned} \quad (4.46)$$

Define the function $y: \mathbb{N} \times \mathbb{N} \times [0, \infty) \rightarrow [0, \infty]$ by

$$\mathbb{N} \times \mathbb{N} \times [0, \infty) \ni (N, n, t) \mapsto y_t^{N,n} := \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_t^N(i)}{(H_t^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_t^N(i))^2} \right) \right]. \quad (4.47)$$

Recall from the beginning of the proof that it is assumed for all $N, n \in \mathbb{N}$ that $y_0^{N,n} < \infty$. Now, applying Itô's lemma and using (4.46), it holds for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ that

$$\begin{aligned} &\mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_{t \wedge \tau_l^{N,n}}^N(i)}{(H_{t \wedge \tau_l^{N,n}}^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_{t \wedge \tau_l^{N,n}}^N(i))^2} \right) \right] - y_0^{N,n} \\ &= \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} 2c_1 \frac{1}{(H_u^N(i))^2} \left(\kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) - (\kappa_P^N + \nu) P_u^N(i) \right. \right. \\ &\quad \left. \left. - \gamma (P_u^N(i))^2 + (\eta - \rho F_u^N(i)) P_u^N(i) H_u^N(i) + \iota_P^N \right) - \left(2c_1 \frac{2P_u^N(i)}{(H_u^N(i))^3} \right. \right. \\ &\quad \left. \left. + \frac{1}{2\delta} \frac{2}{(H_u^N(i))^3} \right) \left(\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + (-\kappa_H^N + \lambda - \alpha^N F_u^N(i)) H_u^N(i) \right. \right. \\ &\quad \left. \left. - \frac{\lambda}{K} (H_u^N(i))^2 - \delta P_u^N(i) H_u^N(i) + \iota_H^N \right) + \frac{1}{2} 2c_1 \frac{6P_u^N(i)}{(H_u^N(i))^4} \beta_H^N H_u^N(i) \right. \\ &\quad \left. \left. + \frac{1}{2} \frac{1}{2\delta} \frac{6}{(H_u^N(i))^4} \beta_H^N H_u^N(i) du \right] \right]. \end{aligned} \quad (4.48)$$

Dropping some negative terms shows for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_u^N(i)}{(H_u^N(i))^{t \wedge \tau_l^{N,n}}} + \frac{1}{2\delta} \frac{1}{(H_u^N(i))^{t \wedge \tau_l^{N,n}}} \right) \right] - y_0^{N,n} \\
 & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \frac{2c_1}{(H_u^N(i))^2} \left(\kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) - \nu P_u^N(i) - \gamma (P_u^N(i))^2 \right. \right. \\
 & \quad \left. \left. + \eta P_u^N(i) H_u^N(i) + \iota_P^N \right) - \left(4c_1 \frac{P_u^N(i)}{(H_u^N(i))^3} + \frac{1}{\delta} \frac{1}{(H_u^N(i))^3} \right) \left((-\kappa_H^N + \lambda - \alpha^N) H_u^N(i) \right. \right. \\
 & \quad \left. \left. - \frac{\lambda}{K} (H_u^N(i))^2 - \delta P_u^N(i) H_u^N(i) + \iota_H^N \right) + 6c_1 \frac{P_u^N(i)}{(H_u^N(i))^3} \beta_H^N + \frac{3}{2\delta} \frac{1}{(H_u^N(i))^3} \beta_H^N du \right] \quad (4.49) \\
 & = \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} 2c_1 \left(\kappa_P^N \frac{1}{(H_u^N(i))^2} \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) - \nu \frac{P_u^N(i)}{(H_u^N(i))^2} - \gamma \frac{(P_u^N(i))^2}{(H_u^N(i))^2} \right. \right. \\
 & \quad \left. \left. + \eta \frac{P_u^N(i)}{H_u^N(i)} + \iota_P^N \frac{1}{(H_u^N(i))^2} - 2(-\kappa_H^N + \lambda - \alpha^N) \frac{P_u^N(i)}{(H_u^N(i))^2} + 2\frac{\lambda}{K} \frac{P_u^N(i)}{H_u^N(i)} + 2\delta \frac{(P_u^N(i))^2}{(H_u^N(i))^2} \right. \right. \\
 & \quad \left. \left. - 2\iota_H^N \frac{P_u^N(i)}{(H_u^N(i))^3} + 3\frac{P_u^N(i)}{(H_u^N(i))^3} \beta_H^N \right) + \frac{\kappa_H^N - \lambda + \alpha^N}{\delta} \frac{1}{(H_u^N(i))^2} + \frac{\lambda}{K\delta} \frac{1}{H_u^N(i)} + \frac{P_u^N(i)}{(H_u^N(i))^2} - \frac{\iota_H^N}{\delta} \frac{1}{(H_u^N(i))^3} \right. \\
 & \quad \left. + \frac{3\beta_H^N}{2\delta} \frac{1}{(H_u^N(i))^3} du \right].
 \end{aligned}$$

Using Young's inequality as well as Lemma 4.1.5 gives for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$

that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_u^N(i))^2} \right) \right] - y_0^{N,n} \\
 & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} 2c_1 \left(\frac{2}{3} \kappa_P^N \frac{1}{(H_u^N(i))^3} + \frac{1}{3} \kappa_P^N c (P_u^N(i))^3 - \nu \frac{P_u^N(i)}{(H_u^N(i))^2} \right. \right. \\
 & \quad - \gamma \frac{(P_u^N(i))^2}{(H_u^N(i))^2} + \frac{1}{2} \frac{\lambda}{2\eta} \eta \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2} \frac{2\eta}{\lambda} \eta P_u^N(i) + \iota_P^N \frac{1}{(H_u^N(i))^2} \\
 & \quad - 2(-\kappa_H^N + \lambda - \alpha^N) \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2} \frac{K}{4} 2 \frac{\lambda}{K} \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2} \frac{4}{K} 2 \frac{\lambda}{K} P_u^N(i) + 2\delta \frac{(P_u^N(i))^2}{(H_u^N(i))^2} \\
 & \quad \left. \left. - 2\iota_H^N \frac{P_u^N(i)}{(H_u^N(i))^3} + 3 \frac{P_u^N(i)}{(H_u^N(i))^3} \beta_H^N \right) + \frac{\kappa_H^N - \lambda + \alpha^N}{\delta} \frac{1}{(H_u^N(i))^2} + \frac{1}{2} \frac{K}{2} \frac{\lambda}{K\delta} \frac{1}{(H_u^N(i))^2} + \frac{1}{2} \frac{2}{K} \frac{\lambda}{K\delta} \right. \\
 & \quad \left. + \frac{P_u^N(i)}{(H_u^N(i))^2} - \frac{\iota_H^N}{\delta} \frac{1}{(H_u^N(i))^3} + \frac{3\beta_H^N}{2\delta} \frac{1}{(H_u^N(i))^3} du \right] \\
 & = \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left[\frac{4c_1}{3} \kappa_P^N - \frac{1}{\delta} \iota_H^N + \frac{3}{2\delta} \beta_H^N \right] \frac{1}{(H_u^N(i))^3} + \frac{2c_1 \kappa_P^N c}{3} (P_u^N(i))^3 \right. \\
 & \quad + \left[c_1 \left(-2\nu + \frac{\lambda}{2} + 4(\kappa_H^N - \lambda + \alpha^N) + \frac{\lambda}{2} \right) + 1 \right] \frac{P_u^N(i)}{(H_u^N(i))^2} + 2c_1 [-\gamma + 2\delta] \frac{(P_u^N(i))^2}{(H_u^N(i))^2} \\
 & \quad + 2c_1 \left[\frac{\eta^2}{\lambda} + \frac{4\lambda}{K^2} \right] P_u^N(i) + \left[2c_1 \iota_P^N + \frac{\kappa_H^N - \lambda + \alpha^N}{\delta} + \frac{\lambda}{4\delta} \right] \frac{1}{(H_u^N(i))^2} \\
 & \quad \left. + 2c_1 [-2\iota_H^N + 3\beta_H^N] \frac{P_u^N(i)}{(H_u^N(i))^3} + \frac{\lambda}{K^2\delta} du \right].
 \end{aligned} \tag{4.50}$$

Recall $\bar{\kappa}_P = \sup_{N \in \mathbb{N}} \kappa_P^N$ and that for all $N \in \mathbb{N}$ it holds that $\alpha^N + \kappa_H^N \leq \frac{\lambda}{4}$, $\iota_P^N \leq \frac{\lambda(\nu+\lambda)}{8\delta}$, and $\iota_H^N \geq \frac{4\delta\kappa_P^N}{3(\nu+\lambda)} + \frac{3\beta_H^N}{2}$. Furthermore, note that $\frac{\lambda}{2} \leq \frac{1}{2c_1}$. Together with the assumption that $\gamma \geq 2\delta$ it follows for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_u^N(i))^2} \right) \right] - y_0^{N,n} \\
 & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \frac{2c_1}{3} \bar{\kappa}_P c (P_u^N(i))^3 - \frac{P_u^N(i)}{(H_u^N(i))^2} + 2c_1 \left[\frac{\eta^2}{\lambda} + \frac{4\lambda}{K^2} \right] P_u^N(i) \right. \\
 & \quad \left. - \frac{\lambda}{4\delta} \frac{1}{(H_u^N(i))^2} + \frac{\lambda}{K^2\delta} du \right] \\
 & \leq \int_0^t c_0^n du - \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \frac{\lambda}{2} \left(2c_1 \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_u^N(i))^2} \right) du \right].
 \end{aligned} \tag{4.51}$$

Using Tonelli's theorem, Fatou's lemma, and (4.43) this implies for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that

$$\begin{aligned}
 y_t^{N,n} + \int_0^t \frac{\lambda}{2} y_u^{N,n} du &= y_t^{N,n} + \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t \frac{\lambda}{2} \left(2c_1 \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_u^N(i))^2} \right) du \right] \\
 &\leq \liminf_{l \rightarrow \infty} \left(\mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_{t \wedge \tau_l^{N,n}}^N(i)}{(H_{t \wedge \tau_l^{N,n}}^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_{t \wedge \tau_l^{N,n}}^N(i))^2} \right) \right] \right. \\
 &\quad \left. + \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \frac{\lambda}{2} \left(2c_1 \frac{P_u^N(i)}{(H_u^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_u^N(i))^2} \right) du \right] \right) \leq y_0^{N,n} + \int_0^t c_0^n du.
 \end{aligned} \tag{4.52}$$

For every $N, n \in \mathbb{N}$ let $z^{N,n}: [0, \infty) \rightarrow \mathbb{R}$ be a process that for all $t \in [0, \infty)$ satisfies $z_t^{N,n} = z_0^{N,n} + \int_0^t (c_0^n - \frac{\lambda}{2} z_s^{N,n}) ds$ with $z_0^{N,n} = y_0^{N,n}$, where uniqueness follows from local Lipschitz continuity. Due to classical comparison results of the theory of ODEs, the above computation yields for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that $y_t^{N,n} \leq z_t^{N,n}$ and for all $N, n \in \mathbb{N}$ that $\sup_{t \in [0, \infty)} z_t^{N,n} = \max \left\{ z_0^{N,n}, \frac{2c_0^n}{\lambda} \right\}$. For all $n \in \mathbb{N}$ it holds that

$$\begin{aligned}
 \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} y_t^{N,n} &\leq \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} z_t^{N,n} = \max \left\{ \sup_{N \in \mathbb{N}} z_0^{N,n}, \frac{2c_0^n}{\lambda} \right\} \\
 &\leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_0^N}{(H_0^N)^2} + \frac{1}{2\delta} \frac{1}{(H_0^N)^2} \right) \right] + \frac{2c_0^n}{\lambda}.
 \end{aligned} \tag{4.53}$$

Using monotone convergence leads to

$$\begin{aligned}
 \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(\frac{2}{\lambda + \nu} \frac{P_t^N(i)}{(H_t^N(i))^2} + \frac{1}{2\delta} \frac{1}{(H_t^N(i))^2} \right) \right] &\leq \lim_{n \rightarrow \infty} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} y_t^{N,n} \\
 &\leq \lim_{n \rightarrow \infty} \left(\sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(2c_1 \frac{P_0^N}{(H_0^N)^2} + \frac{1}{2\delta} \frac{1}{(H_0^N)^2} \right) \right] + \frac{2c_0^n}{\lambda} \right) \\
 &\leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(\frac{2}{\lambda + \nu} \frac{P_0^N}{(H_0^N)^2} + \frac{1}{2\delta} \frac{1}{(H_0^N)^2} \right) \right] \\
 &\quad + \frac{4\bar{\kappa}_{PC}}{3\lambda(\lambda + \nu)} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i (P_t^N(i))^3 \right] \\
 &\quad + \frac{4}{\lambda(\lambda + \nu)} \left(\frac{\eta^2}{\lambda} + \frac{4\lambda}{K^2} \right) \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i P_t^N(i) \right] + \frac{2}{K^2\delta},
 \end{aligned} \tag{4.54}$$

finishing the proof. \square

Lemma 4.1.8. *Assume the setting of Section 4.1.4.1 and assume $\lambda > \nu$ and $\eta - \rho > \frac{\lambda}{K}$. Furthermore, assume for all $N \in \mathbb{N}$ that $\iota_H^N \geq \frac{1}{2}\beta_H^N$, $\kappa_P^N + \kappa_H^N + \alpha^N \leq \frac{\lambda-\nu}{2}$, $\iota_P^N \geq \beta_P^N$, and $\iota_H^N \geq \beta_H^N$. Let $\hat{\mathcal{D}} \subseteq \mathcal{D}$ be a set. Then it holds that*

$$\begin{aligned} & \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(\frac{(\eta-\rho)-\frac{\lambda}{K}}{2(\bar{\kappa}_P+\nu)} \frac{1}{P_t^N(i)} + \frac{1}{P_t^N(i)H_t^N(i)} \right) \right] \\ & \leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(\frac{(\eta-\rho)-\frac{\lambda}{K}}{2(\bar{\kappa}_P+\nu)} \frac{1}{P_0^N(i)} + \frac{1}{P_0^N(i)H_0^N(i)} \right) \right] + \frac{1}{\min\{\bar{\kappa}_P+\nu, \frac{\lambda-\nu}{2}\}} \left(\gamma \frac{(\eta-\rho)-\frac{\lambda}{K}}{2(\bar{\kappa}_P+\nu)} \right. \\ & \quad \left. + (\gamma + \delta) \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \frac{1}{H_t^N(i)} \right] \right). \end{aligned} \quad (4.55)$$

Proof. If the right-hand side of (4.55) is infinite, then the claim trivially holds. For the remainder of the proof assume the right-hand side of (4.55) to be finite. Define $\mathcal{D}_0 := \emptyset$ and for every $n \in \mathbb{N}$ let $\mathcal{D}_n \subseteq \hat{\mathcal{D}}$ be a set with $|\mathcal{D}_n| = \min\{n, |\hat{\mathcal{D}}|\}$ and $\mathcal{D}_n \supseteq \mathcal{D}_{n-1}$. Define $c_0 := \frac{1}{2(\bar{\kappa}_P+\nu)} [(\eta - \rho) - \frac{\lambda}{K}]$ and for every $n \in \mathbb{N}$ let

$$C^n := \gamma c_0 + \left[\gamma + \delta \right] \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \frac{1}{H_t^N(i)} \right]. \quad (4.56)$$

Note that the assumption $\eta - \rho > \frac{\lambda}{K}$ implies $c_0 \in (0, \infty)$. For all $N, n, l \in \mathbb{N}$ define $[0, \infty]$ -valued stopping times

$$\tau_l^{N,n} := \inf \left(\left\{ t \in [0, \infty) : \sum_{i \in \mathcal{D}_n} \sigma_i \left((P_t^N(i))^{-1} + (H_t^N(i))^{-1} \right) > l \right\} \cup \infty \right). \quad (4.57)$$

Lemmas 4.1.3 and 4.1.4 imply that for all $N, n \in \mathbb{N}$ the map

$$[0, \infty) \ni t \mapsto \sum_{i \in \mathcal{D}_n} \sigma_i \left((P_t^N(i))^{-1} + (H_t^N(i))^{-1} \right) \in \mathbb{R} \quad (4.58)$$

is \mathbb{P} -a.s. continuous which implies for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that

$$\mathbb{P} \left[\lim_{l \rightarrow \infty} \tau_l^{N,n} < t \right] = 0. \quad (4.59)$$

For all $t \in [0, \infty)$, $N, n, l \in \mathbb{N}$ Young's inequality leads to

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left(\frac{\sqrt{\beta_P^N P_u^N(i)}}{(P_u^N(i))^2} \left(c_0 + \frac{1}{H_u^N(i)} \right) \right)^2 du \right] \\
 & \leq \bar{\beta}_P \mathbb{E} \left[t \sup_{u \in [0, t]} \sum_{i \in \mathcal{D}_n} \frac{\sigma_i^5}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} \left(\frac{3}{5} \left(P_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-5} + \frac{2}{5} \left(c_0 \right. \right. \right. \\
 & \quad \left. \left. \left. + \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} \right)^5 \right) du \right] \leq \frac{\bar{\beta}_P}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} \mathbb{E} \left[t \sup_{u \in [0, t]} \left(\sum_{i \in \mathcal{D}_n} \sigma_i \left(\left(P_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} \right. \right. \right. \\
 & \quad \left. \left. \left. + \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} + c_0 \right)^5 \right) \right] \leq \frac{\bar{\beta}_P t (l + c_0 \|\mathbf{1}\|_\sigma)^5}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} < \infty
 \end{aligned} \tag{4.60}$$

and

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left(\frac{\sqrt{\beta_H^N H_u^N(i)}}{(H_u^N(i))^2 P_u^N(i)} \right)^2 du \right] \\
 & \leq \bar{\beta}_H \mathbb{E} \left[t \sup_{u \in [0, t]} \sum_{i \in \mathcal{D}_n} \frac{\sigma_i^5}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} \left(\frac{3}{5} \left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-5} + \frac{2}{5} \left(P_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-5} \right) du \right] \\
 & \leq t \sup_{u \in [0, t]} \frac{\bar{\beta}_H}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} \mathbb{E} \left[\left(\sum_{i \in \mathcal{D}_n} \sigma_i \left(\left(H_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} + \left(P_{u \wedge \tau_l^{N,n}}^N(i) \right)^{-1} \right) \right)^5 \right] \\
 & \leq \frac{t \bar{\beta}_H l^5}{\min_{k \in \mathcal{D}_n} \{\sigma_k^4\}} < \infty.
 \end{aligned} \tag{4.61}$$

Hence, for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ it holds that

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i \sqrt{\beta_P^N P_t^N(i)} \frac{1}{(P_t^N(i))^2} \left(c_0 + \frac{1}{H_t^N(i)} \right) dW_u^{P,N}(i) \right] = 0, \\
 & \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \sum_{i \in \mathcal{D}_n} \sigma_i \frac{\sqrt{\beta_H^N H_t^N(i)}}{(H_t^N(i))^2 P_t^N(i)} dW_u^{H,N}(i) \right] = 0.
 \end{aligned} \tag{4.62}$$

Define the function $y: \mathbb{N} \times \mathbb{N} \times [0, \infty) \rightarrow [0, \infty]$ by

$$\mathbb{N} \times \mathbb{N} \times [0, \infty) \ni (N, n, t) \mapsto y_t^{N,n} := \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_t^N(i)} + \frac{1}{P_t^N(i) H_t^N(i)} \right) \right]. \tag{4.63}$$

Recall from the beginning of the proof that it is assumed for all $N, n \in \mathbb{N}$ that $y_0^{N,n} < \infty$.

Applying Itô's lemma and using (4.62), implies for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_{t \wedge \tau_l^{N,n}}^N(i)} + \frac{1}{P_{t \wedge \tau_l^{N,n}}^N(i) H_{t \wedge \tau_l^{N,n}}^N(i)} \right) \right] - y_0^{N,n} \\
 &= \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} - \left(c_0 \frac{1}{(P_u^N(i))^2} + \frac{1}{(P_u^N(i))^2 H_u^N(i)} \right) \left(\kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) \right. \right. \\
 &\quad \left. \left. - (\kappa_P^N + \nu) P_u^N(i) - \gamma (P_u^N(i))^2 + (\eta - \rho F_u^N(i)) P_u^N(i) H_u^N(i) + \iota_P^N \right) \right. \\
 &\quad \left. + \frac{1}{2} c_0 \frac{2}{(P_u^N(i))^3} \beta_P^N P_u^N(i) + \frac{1}{2} \frac{2}{(P_u^N(i))^3 H_u^N(i)} \beta_P^N P_u^N(i) \right. \\
 &\quad \left. - \frac{1}{P_u^N(i) (H_u^N(i))^2} \left(\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + (-\kappa_H^N + \lambda - \alpha^N F_u^N(i)) H_u^N(i) \right) \right. \\
 &\quad \left. - \frac{\lambda}{K} (H_u^N(i))^2 - \delta P_u^N(i) H_u^N(i) + \iota_H^N \right) + \frac{1}{2} \frac{2}{P_u^N(i) (H_u^N(i))^3} \beta_H^N H_u^N(i) du \Big]. \tag{4.64}
 \end{aligned}$$

Dropping some negative terms, now yields for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_u^N(i)} + \frac{1}{P_u^N(i) H_u^N(i)} \right) \right] - y_0^{N,n} \\
 & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} - \left(c_0 \frac{1}{(P_u^N(i))^2} + \frac{1}{(P_u^N(i))^2 H_u^N(i)} \right) \left(-(\kappa_P^N + \nu) P_u^N(i) \right. \right. \\
 & \quad \left. \left. - \gamma (P_u^N(i))^2 + (\eta - \rho) P_u^N(i) H_u^N(i) + \iota_P^N \right) + c_0 \beta_P^N \frac{1}{(P_u^N(i))^2} + \beta_P^N \frac{1}{(P_u^N(i))^2 H_u^N(i)} \right. \\
 & \quad \left. + \beta_H^N \frac{1}{P_u^N(i) (H_u^N(i))^2} - \frac{1}{P_u^N(i) (H_u^N(i))^2} \left((-\kappa_H^N + \lambda - \alpha^N) H_u^N(i) \right. \right. \\
 & \quad \left. \left. - \frac{\lambda}{K} (H_u^N(i))^2 - \delta P_u^N(i) H_u^N(i) + \iota_H^N \right) du \right] \\
 & = \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} (\kappa_P^N + \nu) c_0 \frac{1}{P_u^N(i)} + \gamma c_0 - (\eta - \rho) c_0 \frac{H_u^N(i)}{P_u^N(i)} - \iota_P^N c_0 \frac{1}{(P_u^N(i))^2} \right. \\
 & \quad \left. + (\kappa_P^N + \nu) \frac{1}{P_u^N(i) H_u^N(i)} + \gamma \frac{1}{H_u^N(i)} - (\eta - \rho) \frac{1}{P_u^N(i)} - \iota_P^N \frac{1}{(P_u^N(i))^2 H_u^N(i)} \right. \\
 & \quad \left. + c_0 \beta_P^N \frac{1}{(P_u^N(i))^2} + \beta_P^N \frac{1}{(P_u^N(i))^2 H_u^N(i)} + \beta_H^N \frac{1}{P_u^N(i) (H_u^N(i))^2} - (-\kappa_H^N + \lambda \right. \\
 & \quad \left. - \alpha^N) \frac{1}{P_u^N(i) H_u^N(i)} + \frac{\lambda}{K} \frac{1}{P_u^N(i)} + \delta \frac{1}{H_u^N(i)} - \iota_H^N \frac{1}{P_u^N(i) (H_u^N(i))^2} du \right] \\
 & = \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} \left[(\kappa_P^N + \nu) c_0 - (\eta - \rho) + \frac{\lambda}{K} \right] \frac{1}{P_u^N(i)} + \gamma c_0 - (\eta - \rho) c_0 \frac{H_u^N(i)}{P_u^N(i)} \right. \\
 & \quad \left. + [-\iota_P^N c_0 + c_0 \beta_P^N] \frac{1}{(P_u^N(i))^2} + [(\kappa_P^N + \nu) - (-\kappa_H^N + \lambda - \alpha^N)] \frac{1}{P_u^N(i) H_u^N(i)} \right. \\
 & \quad \left. + [\gamma + \delta] \frac{1}{H_u^N(i)} + [-\iota_P^N + \beta_P^N] \frac{1}{(P_u^N(i))^2 H_u^N(i)} + [\beta_H^N - \iota_H^N] \frac{1}{P_u^N(i) (H_u^N(i))^2} du \right].
 \end{aligned} \tag{4.65}$$

Recall from Section 4.1.4.1 that $\bar{\kappa}_P = \sup_{N \in \mathbb{N}} \kappa_P^N$, and from Assumption 4.1.1 that $\lambda > \nu$, $\eta - \rho > \frac{\lambda}{K}$ and that for all $N \in \mathbb{N}$ it holds that $\kappa_P^N + \kappa_H^N + \alpha^N \leq \frac{\lambda - \nu}{2}$, $\iota_P^N \geq \beta_P^N$, and

$\iota_H^N \geq \beta_H^N$. Hence, for all $t \in [0, \infty)$ and all $N, n, l \in \mathbb{N}$ it follows that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_{t \wedge \tau_l^{N,n}}^N(i)} + \frac{1}{P_{t \wedge \tau_l^{N,n}}^N(i) H_{t \wedge \tau_l^{N,n}}^N(i)} \right) \right] - y_0^{N,n} \\
 & \leq \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^{t \wedge \tau_l^{N,n}} -(\bar{\kappa}_P + \nu) c_0 \frac{1}{P_u^N(i)} + \gamma c_0 - \frac{\lambda - \nu}{2} \frac{1}{P_u^N(i) H_u^N(i)} + \left[\gamma \right. \right. \\
 & \quad \left. \left. + \delta \right] \frac{1}{H_u^N(i)} du \right] \leq \int_0^t C^n du - \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \min \left\{ \bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2} \right\} \sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_u^N(i)} \right. \right. \\
 & \quad \left. \left. + \frac{1}{P_u^N(i) H_u^N(i)} \right) du \right]. \tag{4.66}
 \end{aligned}$$

Applying Tonelli's theorem, Fatou's lemma, and (4.59) for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ it follows that

$$\begin{aligned}
 & y_t^{N,n} + \int_0^t \min \left\{ \bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2} \right\} y_u^{N,n} du \\
 & = y_t^{N,n} + \mathbb{E} \left[\int_0^t \min \left\{ \bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2} \right\} \sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_u^N(i)} + \frac{1}{P_u^N(i) H_u^N(i)} \right) du \right] \\
 & \leq \liminf_{l \rightarrow \infty} \left(\mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_{t \wedge \tau_l^{N,n}}^N(i)} + \frac{1}{P_{t \wedge \tau_l^{N,n}}^N(i) H_{t \wedge \tau_l^{N,n}}^N(i)} \right) \right] \right. \\
 & \quad \left. + \mathbb{E} \left[\int_0^{t \wedge \tau_l^{N,n}} \min \left\{ \bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2} \right\} \sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_u^N(i)} + \frac{1}{P_u^N(i) H_u^N(i)} \right) du \right] \right) \\
 & \leq y_0^{N,n} + \int_0^t C^n du. \tag{4.67}
 \end{aligned}$$

For every $N, n \in \mathbb{N}$, let $z^{N,n}: [0, \infty) \rightarrow \mathbb{R}$ be a process that for all $t \in [0, \infty)$ satisfies $z_t^{N,n} = z_0^{N,n} + \int_0^t (C^n - \min \{ \bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2} \} z_s^{N,n}) ds$, with $z_0^{N,n} = y_0^{N,n}$, where uniqueness follows from local Lipschitz continuity. Using classical comparison results from the theory of ODEs, the above computation yields for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that $y_t^{N,n} \leq z_t^{N,n}$ and for all $N, n \in \mathbb{N}$ that $\sup_{t \in [0, \infty)} z_t^{N,n} = \max \left\{ z_0^{N,n}, \frac{C^n}{\min \{ \bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2} \}} \right\}$. Hence, for every

$n \in \mathbb{N}$ it holds that

$$\begin{aligned}
 & \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_t^N(i)} + \frac{1}{P_t^N(i) H_t^N(i)} \right) \right] = \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} y_t^{N,n} \\
 & \leq \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} z_t^{N,n} = \max \left\{ \sup_{N \in \mathbb{N}} z_0^{N,n}, \frac{C^n}{\min\{\bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2}\}} \right\} \\
 & \leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_0^N(i)} + \frac{1}{P_0^N(i) H_0^N(i)} \right) \right] + \frac{C^n}{\min\{\bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2}\}}.
 \end{aligned} \tag{4.68}$$

Monotone convergence now implies that

$$\begin{aligned}
 & \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(c_0 \frac{1}{P_t^N(i)} + \frac{1}{P_t^N(i) H_t^N(i)} \right) \right] \\
 & = \lim_{n \rightarrow \infty} \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_t^N(i)} + \frac{1}{P_t^N(i) H_t^N(i)} \right) \right] \\
 & \leq \lim_{n \rightarrow \infty} \left(\sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \left(c_0 \frac{1}{P_0^N(i)} + \frac{1}{P_0^N(i) H_0^N(i)} \right) \right] + \frac{C^n}{\min\{\bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2}\}} \right) \\
 & = \sup_{N \in \mathbb{N}} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(c_0 \frac{1}{P_0^N(i)} + \frac{1}{P_0^N(i) H_0^N(i)} \right) \right] + \frac{\gamma c_0 + (\gamma + \delta) \sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \frac{1}{H_t^N(i)} \right]}{\min\{\bar{\kappa}_P + \nu, \frac{\lambda - \nu}{2}\}}.
 \end{aligned} \tag{4.69}$$

This completes the proof. \square

Theorem 4.1.9. *Assume the setting of Section 4.1.4.1 and let Assumption 4.1.1 hold. Then for all $(x, y, z) \in (0, \infty)^2 \times [0, 1]$ it holds that*

$$\begin{aligned}
 u(x, y, z) &:= (\eta - \rho z) \left(x - h_\infty(z) - h_\infty(z) \ln \left(\frac{x}{h_\infty(z)} \right) \right) \\
 &+ \delta \left(y - p_\infty(z) - p_\infty(z) \ln \left(\frac{y}{p_\infty(z)} \right) \right) \geq 0.
 \end{aligned} \tag{4.70}$$

Furthermore, there exists a constant $c_0 \in (0, \infty)$ such that for every set $\hat{\mathcal{D}} \subseteq \mathcal{D}$, for every

$N \in \mathbb{N}$, and every $t \in [0, \infty)$ it holds that

$$\begin{aligned} & \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i u \left(H_t^N(i), P_t^N(i), F_t^N(i) \right) \right] + \int_0^t (\eta - \rho) \frac{\lambda}{K} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(H_u^N(i) \right. \right. \\ & \quad \left. \left. - h_\infty \left(F_u^N(i) \right) \right)^2 \right] + \delta \gamma \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i \left(P_u^N(i) - p_\infty \left(F_u^N(i) \right) \right)^2 \right] du \\ & \leq \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i u \left(H_0^N(i), P_0^N(i), F_0^N(i) \right) \right] + t c_0 \max \left\{ \kappa_H^N, \kappa_P^N, \alpha^N, \iota_H^N, \iota_P^N, \beta_H^N, \beta_P^N \right\}. \end{aligned} \quad (4.71)$$

Proof. For the remainder of the proof fix a set $\hat{\mathcal{D}} \subseteq \mathcal{D}$. Define $\mathcal{D}_0 := \emptyset$ and for every $n \in \mathbb{N}$ let $\mathcal{D}_n \subseteq \hat{\mathcal{D}}$ be a set with $|\mathcal{D}_n| = \min \left\{ n, |\hat{\mathcal{D}}| \right\}$ and $\mathcal{D}_n \supseteq \mathcal{D}_{n-1}$. First it will be shown that for all $(x, y, z) \in (0, \infty)^2 \times [0, 1]$ it holds that $u(x, y, z) \geq 0$. Define for all $x \in (0, \infty)$ the real-valued function $(0, \infty) \ni y \mapsto f_x(y) := x - y - y \ln \left(\frac{x}{y} \right)$. For all $x \in (0, \infty)$ the function f_x has for all $y \in (0, \infty)$ first and second order derivatives $\frac{df_x}{dy}(y) = \ln(y) - \ln(x)$ and $\frac{d^2 f_x}{dy^2}(y) = \frac{1}{y} > 0$. Thus, for all $x \in (0, \infty)$ the function f_x has its global minimum at x with $f_x(x) = 0$. Consequently, for any $(x, y) \in (0, \infty)^2$ it holds that $f_x(y) \geq f_x(x) = 0$. This shows that for all $(x, y, z) \in (0, \infty)^2 \times [0, 1]$ it holds that $u(x, y, z) \geq 0$. In order to prove the second part of the claim, a Lyapunov function can be used, that is defined here analogously to the well-known Lyapunov function in the deterministic setting. Define $D_V := \left(l_\sigma^1 \cap (0, \infty)^{\mathcal{D}} \right) \times \left(l_\sigma^1 \cap (0, \infty)^{\mathcal{D}} \right) \times E_1$. For any subset $\hat{\mathcal{D}}' \subseteq \hat{\mathcal{D}}$ define the function $V_{\hat{\mathcal{D}}'}: D_V \rightarrow [0, \infty]$ for any $(h, p, f) \in D_V$ by

$$V_{\hat{\mathcal{D}}'}((h, p, f)) := \sum_{i \in \hat{\mathcal{D}}'} \sigma_i u(h_i, p_i, f_i). \quad (4.72)$$

Due to the non-negativity of the mapping u , it follows for any $\hat{\mathcal{D}}' \subseteq \hat{\mathcal{D}}$ and any $z \in D_V$ that $V_{\hat{\mathcal{D}}'}(z) \in [0, \infty]$ is well-defined. The fact that for all $x \in (0, \infty)$ it holds that $-\ln(x) \leq \sqrt{\frac{1}{x}} \leq \frac{1}{2} \left(\frac{1}{x} + 1 \right)$ as well as the assumption $\sup_{N \in \mathbb{N}} \mathbb{E} \left[\left\| \left(H_0^N + P_0^N \right)^4 + \frac{1}{(H_0^N)^2} + \frac{P_0^N}{(H_0^N)^2} + \frac{1}{P_0^N} + \frac{1}{P_0^N H_0^N} \right\|_\sigma \right] < \infty$ imply

$$\sup_{N \in \mathbb{N}} \mathbb{E} \left[V_{\mathcal{D}} \left(H_0^N, P_0^N, F_0^N \right) \right] < \infty. \quad (4.73)$$

Now the first and second order partial derivatives, which are needed in the application of Itô's lemma below, are calculated. For all $n \in \mathbb{N}$, $z = (h, p, f) \in D_V$, and $i \in \mathcal{D}_n$ it holds that $\frac{dV_{\mathcal{D}_n}}{dh_i}(z) = \sigma_i (\eta - \rho f_i) \left(1 - \frac{h_\infty(f_i)}{h_i} \right)$, $\frac{d^2 V_{\mathcal{D}_n}}{dh_i^2}(z) = \sigma_i (\eta - \rho f_i) \frac{h_\infty(f_i)}{h_i^2}$, $\frac{dV_{\mathcal{D}_n}}{dp_i}(z) =$

$\sigma_i \delta \left(1 - \frac{p_\infty(f_i)}{p_i}\right)$, and $\frac{d^2 V_{\mathcal{D}_n}}{dp_i^2}(z) = \sigma_i \delta \frac{p_\infty(f_i)}{p_i^2}$ as well as

$$\begin{aligned} \frac{dV_{\mathcal{D}_n}}{df_i}(z) = \sigma_i & \left[-\rho \left(h_i - h_\infty(f_i) - h_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) \right) + (\eta - \rho f_i) \left(-h'_\infty(f_i) \right. \right. \\ & \left. \left. - h'_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) - \frac{(h_\infty(f_i))^2}{h_i} \frac{-h_i}{(h_\infty(f_i))^2} h'_\infty(f_i) \right) + \delta \left(-p'_\infty(f_i) \right. \right. \\ & \left. \left. - p'_\infty(f_i) \ln \left(\frac{p_i}{p_\infty(f_i)} \right) - \frac{(p_\infty(f_i))^2}{p_i} \frac{-p_i}{(p_\infty(f_i))^2} p'_\infty(f_i) \right) \right] = \sigma_i \left[-\rho \left(h_i - h_\infty(f_i) \right. \right. \\ & \left. \left. - h_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) \right) - (\eta - \rho f_i) h'_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) - \delta p'_\infty(f_i) \ln \left(\frac{p_i}{p_\infty(f_i)} \right) \right] \end{aligned} \quad (4.74)$$

and

$$\begin{aligned} \frac{d^2 V_{\mathcal{D}_n}}{df_i^2}(z) = \sigma_i & \left[\rho \left(h'_\infty(f_i) + h'_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) + \frac{(h_\infty(f_i))^2}{h_i} (-1) \frac{h_i}{(h_\infty(f_i))^2} h'_\infty(f_i) \right) \right. \\ & + \rho h'_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) - (\eta - \rho f_i) \left(h''_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) \right. \\ & + h'_\infty(f_i) \frac{h_\infty(f_i)}{h_i} \frac{-h_i}{(h_\infty(f_i))^2} h'_\infty(f_i) \left. \right) - \delta \left(p''_\infty(f_i) \ln \left(\frac{p_i}{p_\infty(f_i)} \right) \right. \\ & \left. + p'_\infty(f_i) \frac{p_\infty(f_i)}{p_i} \frac{-p_i}{(p_\infty(f_i))^2} p'_\infty(f_i) \right) \right] \quad (4.75) \\ = \sigma_i & \left[2\rho h'_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) - (\eta - \rho f_i) \left(h''_\infty(f_i) \ln \left(\frac{h_i}{h_\infty(f_i)} \right) \right. \right. \\ & \left. \left. - (h'_\infty(f_i))^2 \frac{1}{h_\infty(f_i)} \right) - \delta \left(p''_\infty(f_i) \ln \left(\frac{p_i}{p_\infty(f_i)} \right) - (p'_\infty(f_i))^2 \frac{1}{p_\infty(f_i)} \right) \right]. \end{aligned}$$

Recall that for all $x \in [0, 1]$ it holds that $h_\infty(x) = \frac{1}{b(a-x)}$ and $p_\infty(x) = \frac{\lambda}{\delta} \left(1 - \frac{1}{Kb(a-x)}\right)$ and note that the assumption that $\eta - \rho > \frac{\nu}{K}$ implies for all $x \in [0, 1]$ that $p_\infty(x) > 0$. Therefore, for all $x \in [0, 1]$ it follows that

$$\begin{aligned} h'_\infty(x) &= \frac{1}{b(a-x)^2} > 0, \quad h''_\infty(x) = \frac{2}{b(a-x)^3} > 0, \\ p'_\infty(x) &= -\frac{\lambda}{\delta Kb(a-x)^2} < 0, \quad p''_\infty(x) = -\frac{2\lambda}{\delta Kb(a-x)^3} < 0. \end{aligned} \quad (4.76)$$

So h_∞ , h'_∞ , and h''_∞ are strictly monotonically increasing on $[0, 1]$ while p_∞ , p'_∞ , and p''_∞ are strictly monotonically decreasing on $[0, 1]$. Also, it holds that $\max_{x \in [0, 1]} \delta p_\infty(x) \leq \lambda$. Observe that for all $x \in (0, \infty)$ it holds that $|\ln(x)| \leq \sqrt{x} + \frac{1}{\sqrt{x}}$. Together with Young's inequality as

well as Lemmas 4.1.6, 4.1.7, and 4.1.8 this implies for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t \left(\sqrt{\beta_P^N P_u^N(i)} \delta \left(1 - \frac{p_\infty(F_u^N(i))}{P_u^N(i)} \right) \right)^2 du \right] \\
 & \leq \bar{\beta}_P \delta^2 \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t P_u^N(i) \left(1 + \frac{(p_\infty(0))^2}{(P_u^N(i))^2} \right) du \right] \\
 & \leq \bar{\beta}_P \delta^2 \sup_{u \in [0, t]} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i t \left(P_u^N(i) + \frac{(p_\infty(0))^2}{P_u^N(i)} \right) \right] < \infty
 \end{aligned} \tag{4.77}$$

and

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t \left(\sqrt{\beta_H^N H_u^N(i)} (\eta - \rho F_u^N(i)) \left(1 - \frac{h_\infty(F_u^N(i))}{H_u^N(i)} \right) \right)^2 du \right] \\
 & \leq \bar{\beta}_H \eta^2 \sup_{u \in [0, t]} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i t \left(H_u^N(i) + \frac{(h_\infty(1))^2}{H_u^N(i)} \right) \right] < \infty
 \end{aligned} \tag{4.78}$$

and

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t \left(\sqrt{\frac{\beta_H^N F_u^N(i)(1-F_u^N(i))}{H_u^N(i)}} \left(-\rho(H_u^N(i) - h_\infty(F_u^N(i))) \right. \right. \right. \\
 & \quad \left. \left. - h_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) \right) - (\eta - \rho F_u^N(i)) h'_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) \right. \\
 & \quad \left. \left. - \delta p'_\infty(F_u^N(i)) \ln \left(\frac{P_u^N(i)}{p_\infty(F_u^N(i))} \right) \right) \right)^2 du \right] \\
 & \leq \bar{\beta}_H \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i \int_0^t \frac{1}{H_u^N(i)} \left(\rho H_u^N(i) + \rho h_\infty(1) + \rho h_\infty(1) \left(\frac{\sqrt{H_u^N(i)}}{\sqrt{h_\infty(0)}} + \frac{\sqrt{h_\infty(1)}}{\sqrt{H_u^N(i)}} \right) \right. \right. \\
 & \quad \left. \left. + \eta h'_\infty(1) \left(\frac{\sqrt{H_u^N(i)}}{\sqrt{h_\infty(0)}} + \frac{\sqrt{h_\infty(1)}}{\sqrt{H_u^N(i)}} \right) + \delta |p'_\infty(1)| \left(\frac{\sqrt{P_u^N(i)}}{\sqrt{p_\infty(1)}} + \frac{\sqrt{p_\infty(0)}}{\sqrt{P_u^N(i)}} \right) \right)^2 du \right] \\
 & \leq \bar{\beta}_H \sup_{u \in [0, t]} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i 2^7 \left(\rho^2 H_u^N(i) + \rho^2 (h_\infty(1))^2 \left(\frac{1}{H_u^N(i)} + \frac{1}{h_\infty(0)} + \frac{h_\infty(1)}{(H_u^N(i))^2} \right) \right. \right. \\
 & \quad \left. \left. + \eta^2 (h'_\infty(1))^2 \left(\frac{1}{h_\infty(0)} + \frac{h_\infty(1)}{(H_u^N(i))^2} \right) + \delta^2 (p'_\infty(1))^2 \left(\frac{1}{p_\infty(1)} \frac{P_u^N(i)}{H_u^N(i)} + \frac{p_\infty(0)}{P_u^N(i) H_u^N(i)} \right) \right) \right] \\
 & < \infty.
 \end{aligned} \tag{4.79}$$

Hence, for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ it follows that

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i \sqrt{\beta_P^N P_u^N(i)} \delta \left(1 - \frac{p_\infty(F_u^N(i))}{P_u^N(i)} \right) dW_u^{P,N}(i) \right] = 0, \\
 & \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i \sqrt{\beta_H^N H_u^N(i)} (\eta - \rho F_u^N(i)) \left(1 - \frac{h_\infty(F_u^N(i))}{H_u^N(i)} \right) dW_u^{H,N}(i) \right] = 0, \\
 & \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i \sqrt{\frac{\beta_H^N F_u^N(i)(1-F_u^N(i))}{H_u^N(i)}} \left[-\rho \left(H_u^N(i) - h_\infty(F_u^N(i)) \right) \right. \right. \\
 & \quad \left. \left. - h_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) \right) - (\eta - \rho F_u^N(i)) h'_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) \right. \right. \\
 & \quad \left. \left. - \delta p'_\infty(F_u^N(i)) \ln \left(\frac{P_u^N(i)}{p_\infty(F_u^N(i))} \right) \right] dW_u^{F,N}(i) \right] = 0.
 \end{aligned} \tag{4.80}$$

For all $t \in [0, \infty)$, all $N \in \mathbb{N}$, and all $i \in \mathcal{D}$ define

$$\begin{aligned}
 R_t^N(i) &:= \max \left\{ \max \left\{ \eta c, \rho c, \rho, c \eta \frac{h'_\infty(1)}{h_\infty(0)}, \eta \frac{h'_\infty(1)}{h_\infty(0)} \right\} H_t^N(i), \eta h_\infty(1), \eta, \right. \\
 & \quad \max \left\{ \frac{\eta}{2} h_\infty(1), \rho (h_\infty(1))^2, \frac{\eta}{2} \frac{(h'_\infty(1))^2}{h_\infty(0)}, \frac{\delta}{2} \frac{(p'_\infty(1))^2}{p_\infty(1)} \right\} \frac{1}{H_t^N(i)}, \delta c P_t^N(i), \delta p_\infty(0), \delta, \\
 & \quad \frac{\delta}{2} \frac{p_\infty(0)}{P_t^N(i)}, \max \left\{ \frac{1}{2} \rho^2 (h_\infty(1))^2, \frac{3}{4} \rho^{\frac{4}{3}} (h_\infty(1))^2, \frac{3}{4} \left(\eta h'_\infty(1) \sqrt{h_\infty(1)} \right)^{\frac{4}{3}}, \frac{1}{4}, \right. \\
 & \quad \left. \frac{\eta}{2} h''_\infty(1) h_\infty(1) \right\} \frac{1}{(H_t^N(i))^2}, \frac{1}{2} c (H_t^N(i))^2, \frac{1}{4} c (H_t^N(i))^4, \frac{1}{2} \left(\frac{\delta |p'_\infty(1)|}{\sqrt{p_\infty(1)}} \right)^2 \frac{P_t^N(i)}{(H_t^N(i))^2}, \\
 & \quad \left. \frac{1}{2} \delta^2 |p'_\infty(1)|^2 \frac{p_\infty(0)}{P_t^N(i) H_t^N(i)}, \delta |p'_\infty(1)| \sqrt{\frac{p_\infty(0)}{P_t^N(i)}}, \rho \frac{h'_\infty(1)}{h_\infty(0)}, \frac{\delta}{2} \frac{|p''_\infty(1)|}{p_\infty(1)} \frac{P_t^N(i)}{H_t^N(i)} \right\}, \\
 b^N &:= \max \left\{ \kappa_H^N, \kappa_P^N, \alpha^N, \iota_H^N, \iota_P^N, \beta_H^N, \beta_P^N \right\}.
 \end{aligned} \tag{4.81}$$

Note that $\lim_{N \rightarrow \infty} b^N = 0$. Define $c_0 := 32 \sup_{M \in \mathbb{N}} \sup_{u \in [0, \infty)} \mathbb{E} [\|R_u^M\|_\sigma]$. Lemmas 4.1.6, 4.1.7, and

4.1.8 imply $c_0 \in (0, \infty)$. For all $t \in [0, \infty)$, all $N \in \mathbb{N}$, and all $a \in \left\{ \eta, \rho, \eta \frac{h'_\infty(1)}{h_\infty(0)} \right\}$ it holds that

$$\sum_{i \in \mathcal{D}} \sigma_i a \sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \leq \sum_{i \in \mathcal{D}} \sigma_i c a H_t^N(i) \leq \sum_{i \in \mathcal{D}} \sigma_i R_t^N(i). \tag{4.82}$$

Furthermore, for all $t \in [0, \infty)$ and all $N \in \mathbb{N}$ it holds that

$$\sum_{i \in \mathcal{D}} \sigma_i \delta \sum_{j \in \mathcal{D}} m(i, j) P_t^N(j) \leq \sum_{i \in \mathcal{D}} \sigma_i \delta c P_t^N(i) \leq \sum_{i \in \mathcal{D}} \sigma_i R_t^N(i). \tag{4.83}$$

Young's inequality and Lemma 4.1.5 yield for all $t \in [0, \infty)$ and all $N \in \mathbb{N}$ that

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}} \sigma_i \rho^{\frac{h_\infty(F_t^N(i))}{H_t^N(i)}} \sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(\frac{1}{2} \left(\rho^{\frac{h_\infty(F_t^N(i))}{H_t^N(i)}} \right)^2 + \frac{1}{2} \left(\sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \right)^2 \right) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(R_t^N(i) + \frac{1}{2} c (H_t^N(i))^2 \right) \leq \sum_{i \in \mathcal{D}} \sigma_i 2R_t^N(i),
 \end{aligned} \tag{4.84}$$

and

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}} \sigma_i \frac{-\delta p'_\infty(F_t^N(i))}{\sqrt{p_\infty(F_t^N(i))}} \frac{\sqrt{P_t^N(i)}}{H_t^N(i)} \sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \frac{1}{2} \left(\left(\frac{\delta |p'_\infty(1)|}{\sqrt{p_\infty(1)}} \frac{\sqrt{P_t^N(i)}}{H_t^N(i)} \right)^2 + \left(\sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \right)^2 \right) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(R_t^N(i) + \frac{1}{2} c (H_t^N(i))^2 \right) \leq \sum_{i \in \mathcal{D}} \sigma_i 2R_t^N(i),
 \end{aligned} \tag{4.85}$$

and

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}} \sigma_i (-1) \delta p'_\infty(F_t^N(i)) \frac{\sqrt{p_\infty(F_t^N(i))}}{\sqrt{P_t^N(i) H_t^N(i)}} \sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(\frac{1}{2} \delta^2 |p'_\infty(1)|^2 \frac{p_\infty(0)}{P_t^N(i) H_t^N(i)} + \frac{1}{2} \left(\sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \right)^2 \frac{1}{H_t^N(i)} \right) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(R_t^N(i) + \frac{1}{4} \left(\sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \right)^4 + \frac{1}{4} \frac{1}{(H_t^N(i))^2} \right) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(R_t^N(i) + \frac{1}{4} c (H_t^N(i))^4 + R_t^N(i) \right) \leq \sum_{i \in \mathcal{D}} \sigma_i 3R_t^N(i).
 \end{aligned} \tag{4.86}$$

Again using Young's inequality and Lemma 4.1.5 it follows that for all $t \in [0, \infty)$, all

$a \in \left\{ \rho(h_\infty(1))^{\frac{3}{2}}, \eta h'_\infty(1) \sqrt{h_\infty(1)} \right\}$, and all $N \in \mathbb{N}$ that

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}} \sigma_i a \left(\frac{1}{H_t^N(i)} \right)^{\frac{3}{2}} \sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(\frac{3}{4} a^{\frac{4}{3}} \left(\frac{1}{H_t^N(i)} \right)^2 + \frac{1}{4} \left(\sum_{j \in \mathcal{D}} m(i, j) H_t^N(j) \right)^4 \right) \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \left(\mathbb{R}_t^N(i) + \frac{1}{4} c (H_t^N(i))^4 \right) \leq \sum_{i \in \mathcal{D}} \sigma_i 2R_t^N(i).
 \end{aligned} \tag{4.87}$$

Due to Lemma 4.1.2 it follows that $W^{H,N}(i)$, $W^{F,N}(i)$, $N \in \mathbb{N}$, $i \in \mathcal{D}$, are independent Brownian motions and due to Lemmas 4.1.6, 4.1.7, and 4.1.8 it holds for all $t \in [0, \infty)$ and all $N \in \mathbb{N}$ that \mathbb{P} -a.s. $(H_t^N, P_t^N, F_t^N) \in D_V$. Thus, Itô's lemma and (4.80) imply for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} [V_{\mathcal{D}_n}((H_t^N, P_t^N, F_t^N))] - \mathbb{E} [V_{\mathcal{D}_n}((H_0^N, P_0^N, F_0^N))] \\
 & = \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i \left((\eta - \rho F_u^N(i)) \left(1 - \frac{h_\infty(F_u^N(i))}{H_u^N(i)} \right) \left\{ \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (H_u^N(j) \right. \right. \right. \\
 & \quad \left. \left. - H_u^N(i)) + H_u^N(i) \left[\lambda \left(1 - \frac{H_u^N(i)}{K} \right) - \delta P_u^N(i) - \alpha^N F_u^N(i) \right] + \iota_H^N \right\} \right. \\
 & \quad \left. + \frac{(\eta - \rho F_u^N(i)) h_\infty(F_u^N(i))}{2 (H_u^N(i))^2} \beta_H^N H_u^N(i) + \delta \left(1 - \frac{p_\infty(F_u^N(i))}{P_u^N(i)} \right) \left\{ \kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) (P_u^N(j) \right. \right. \right. \\
 & \quad \left. \left. - P_u^N(i)) + P_u^N(i) [-\nu - \gamma P_u^N(i) + (\eta - \rho F_u^N(i)) H_u^N(i)] + \iota_P^N \right\} \right. \\
 & \quad \left. + \frac{\delta p_\infty(F_u^N(i))}{2 (P_u^N(i))^2} \beta_P^N P_u^N(i) + \left[-\rho \left(H_u^N(i) - h_\infty(F_u^N(i)) \right) \right. \right. \\
 & \quad \left. \left. - h_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) - (\eta - \rho F_u^N(i)) h'_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) \right. \right. \\
 & \quad \left. \left. - \delta p'_\infty(F_u^N(i)) \ln \left(\frac{P_u^N(i)}{p_\infty(F_u^N(i))} \right) \right] \left\{ \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (F_u^N(j) - F_u^N(i)) \frac{H_u^N(j)}{H_u^N(i)} \right. \right. \\
 & \quad \left. \left. - \alpha^N F_u^N(i) (1 - F_u^N(i)) \right\} + \left\{ \rho h'_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) \right. \right. \\
 & \quad \left. \left. - \frac{\eta - \rho F_u^N(i)}{2} \left(h''_\infty(F_u^N(i)) \ln \left(\frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) - \frac{(h'_\infty(F_u^N(i)))^2}{h_\infty(F_u^N(i))} \right) \right. \right. \\
 & \quad \left. \left. - \frac{\delta}{2} \left(p''_\infty(F_u^N(i)) \ln \left(\frac{P_u^N(i)}{p_\infty(F_u^N(i))} \right) - \frac{(p'_\infty(F_u^N(i)))^2}{p_\infty(F_u^N(i))} \right) \right\} \frac{\beta_H^N F_u^N(i) (1 - F_u^N(i))}{H_u^N(i)} \right) du \right].
 \end{aligned} \tag{4.88}$$

Note that for all $x \in [0, 1]$ it holds that $0 < \eta - \rho x \leq \eta$. Together with the fact that for all

$x \in (0, \infty)$ it holds that $\ln(x) \leq \sqrt{x}$, $\ln(x) \leq x$, $|\ln(x)| \leq \sqrt{x} + \sqrt{\frac{1}{x}}$, and $|\ln(x)| \leq x + \sqrt{\frac{1}{x}}$ and dropping negative terms, this implies for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} [V_{\mathcal{D}_n} ((H_t^N, P_t^N, F_t^N))] - \mathbb{E} [V_{\mathcal{D}_n} ((H_0^N, P_0^N, F_0^N))] \\
 & \leq \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}_n} \sigma_i \left(\eta \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + \eta h_\infty(F_u^N(i)) \kappa_H^N \right. \right. \\
 & \quad + (\eta - \rho F_u^N(i)) (H_u^N(i) - h_\infty(F_u^N(i))) \left[\lambda \left(1 - \frac{H_u^N(i)}{K} \right) - \delta P_u^N(i) \right] \\
 & \quad + \eta h_\infty(F_u^N(i)) \alpha^N + \eta \iota_H^N + \frac{\eta}{2} h_\infty(F_u^N(i)) \beta_H^N \frac{1}{H_u^N(i)} \\
 & \quad + \delta \kappa_P^N \sum_{j \in \mathcal{D}} m(i, j) P_u^N(j) + \delta p_\infty(F_u^N(i)) \kappa_P^N + \delta (P_u^N(i) - p_\infty(F_u^N(i))) [-\nu \\
 & \quad - \gamma P_u^N(i) + (\eta - \rho F_u^N(i)) H_u^N(i)] + \delta \iota_P^N + \frac{\delta}{2} \frac{p_\infty(F_u^N(i))}{P_u^N(i)} \beta_P^N \\
 & \quad + \rho \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + \rho H_u^N(i) \alpha^N + \rho \frac{h_\infty(F_u^N(i))}{H_u^N(i)} \left(1 + \frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right. \\
 & \quad \left. \left. + \sqrt{\frac{h_\infty(F_u^N(i))}{H_u^N(i)}} \right) \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) H_u^N(j) + \rho h_\infty(F_u^N(i)) \frac{h_\infty(F_u^N(i))}{H_u^N(i)} \alpha^N \right. \\
 & \quad + \left[\eta h'_\infty(F_u^N(i)) \left(\sqrt{\frac{h_\infty(F_u^N(i))}{H_u^N(i)}} + \frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right) - \delta p'_\infty(F_u^N(i)) \left(\sqrt{\frac{P_u^N(i)}{p_\infty(F_u^N(i))}} \right. \right. \\
 & \quad \left. \left. + \sqrt{\frac{p_\infty(F_u^N(i))}{P_u^N(i)}} \right) \right] \kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) \frac{H_u^N(j)}{H_u^N(i)} + \left[\eta h'_\infty(F_u^N(i)) \frac{H_u^N(i)}{h_\infty(F_u^N(i))} \right. \\
 & \quad \left. - \delta p'_\infty(F_u^N(i)) \sqrt{\frac{p_\infty(F_u^N(i))}{P_u^N(i)}} \right] \alpha^N + \rho h'_\infty(F_u^N(i)) \frac{H_u^N(i)}{h_\infty(F_u^N(i))} \frac{\beta_H^N}{H_u^N(i)} \\
 & \quad + \frac{\eta}{2} \left(h''_\infty(F_u^N(i)) \frac{h_\infty(F_u^N(i))}{H_u^N(i)} + \frac{(h'_\infty(F_u^N(i)))^2}{h_\infty(F_u^N(i))} \right) \frac{\beta_H^N}{H_u^N(i)} \\
 & \quad \left. + \frac{\delta}{2} \left(-p''_\infty(F_u^N(i)) \frac{P_u^N(i)}{p_\infty(F_u^N(i))} + \frac{(p'_\infty(F_u^N(i)))^2}{p_\infty(F_u^N(i))} \right) \frac{\beta_H^N}{H_u^N(i)} \right] du \right]. \tag{4.89}
 \end{aligned}$$

Using (4.82), (4.83), (4.84), (4.85), (4.86), and (4.87) it follows for all $t \in [0, \infty)$ and all

$N, n \in \mathbb{N}$ that

$$\begin{aligned}
 & \mathbb{E} [V_{\mathcal{D}_n} ((H_t^N, P_t^N, F_t^N))] - \mathbb{E} [V_{\mathcal{D}_n} ((H_0^N, P_0^N, F_0^N))] \\
 & \leq \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}} \sigma_i \left(b^N 32 R_u^N(i) + (\eta - \rho F_u^N(i)) (H_u^N(i) \right. \right. \\
 & \quad \left. \left. - h_\infty(F_u^N(i))) \left[\lambda \left(1 - \frac{H_u^N(i)}{K} \right) - \delta P_u^N(i) \right] + \delta (P_u^N(i) - p_\infty(F_u^N(i))) \left[-\nu \right. \right. \right. \\
 & \quad \left. \left. - \gamma P_u^N(i) + (\eta - \rho F_u^N(i)) H_u^N(i) \right] \right) du \right]. \tag{4.90}
 \end{aligned}$$

Note that for all $x \in [0, 1]$ it holds that

$$\begin{aligned}
 \delta p_\infty(x) + \frac{\lambda}{K} h_\infty(x) - \lambda &= \frac{\delta \lambda K(\eta - \rho x) - \delta \lambda \nu + \lambda \delta \nu + \lambda^2 \gamma}{\lambda \gamma + \delta K(\eta - \rho x)} - \lambda = \frac{\delta K(\eta - \rho x) + \lambda \gamma}{\lambda \gamma + \delta K(\eta - \rho x)} \lambda - \lambda = 0, \\
 \nu - (\eta - \rho x) h_\infty(x) + \gamma p_\infty(x) &= \nu - \frac{(\eta - \rho x) K \delta \nu + (\eta - \rho x) K \gamma \lambda - \gamma \lambda K(\eta - \rho x) + \gamma \lambda \nu}{\lambda \gamma + \delta K(\eta - \rho x)} \\
 &= \nu - \frac{(\eta - \rho x) K \delta + \gamma \lambda}{\lambda \gamma + \delta K(\eta - \rho x)} \nu = 0. \tag{4.91}
 \end{aligned}$$

From (4.91) it follows that for all $t \in [0, \infty)$ and all $N, n \in \mathbb{N}$

$$\begin{aligned}
 & \mathbb{E} [V_{\mathcal{D}_n} ((H_t^N, P_t^N, F_t^N))] - \mathbb{E} [V_{\mathcal{D}_n} ((H_0^N, P_0^N, F_0^N))] \\
 & \leq \mathbb{E} \left[\int_0^t \sum_{i \in \mathcal{D}} \sigma_i \left(b^N 32 R_u^N(i) + (\eta - \rho F_u^N(i)) (H_u^N(i) - h_\infty(F_u^N(i))) \left[\lambda \right. \right. \right. \\
 & \quad \left. \left. - \frac{\lambda}{K} (H_u^N(i) - h_\infty(F_u^N(i))) - \delta (P_u^N(i) - p_\infty(F_u^N(i))) - \lambda \right] \right. \\
 & \quad \left. + \delta (P_u^N(i) - p_\infty(F_u^N(i))) \left[-\nu - \gamma (P_u^N(i) - p_\infty(F_u^N(i))) \right. \right. \\
 & \quad \left. \left. + (\eta - \rho F_u^N(i)) (H_u^N(i) - h_\infty(F_u^N(i))) + \nu \right] \right) du \right]. \tag{4.92}
 \end{aligned}$$

Hence, it follows for every $N, n \in \mathbb{N}$ and every $t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} [V_{\mathcal{D}_n} (H_t^N, P_t^N, F_t^N)] + \int_0^t (\eta - \rho) \frac{\lambda}{K} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (H_u^N(i) - h_\infty(F_u^N(i)))^2 \right] \\
 & \quad + \delta \gamma \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (P_u^N(i) - p_\infty(F_u^N(i)))^2 \right] du \\
 & \leq \mathbb{E} [V_{\mathcal{D}} (H_0^N, P_0^N, F_0^N)] + t b^N 32 \sup_{M \in \mathbb{N}} \sup_{u \in [0, \infty)} \mathbb{E} [\|R_u^M\|_\sigma]. \tag{4.93}
 \end{aligned}$$

Applying monotone convergence now yields for every $N \in \mathbb{N}$ and every $t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} [V_{\hat{\mathcal{D}}} (H_t^N, P_t^N, F_t^N)] + \int_0^t (\eta - \rho) \frac{\lambda}{K} \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i (H_u^N(i) - h_\infty(F_u^N(i)))^2 \right] \\
 & + \delta \gamma \mathbb{E} \left[\sum_{i \in \hat{\mathcal{D}}} \sigma_i (P_u^N(i) - p_\infty(F_u^N(i)))^2 \right] du = \lim_{n \rightarrow \infty} \left(\mathbb{E} [V_{\mathcal{D}_n} (H_t^N, P_t^N, F_t^N)] \right. \\
 & + \int_0^t (\eta - \rho) \frac{\lambda}{K} \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (H_u^N(i) - h_\infty(F_u^N(i)))^2 \right] + \delta \gamma \mathbb{E} \left[\sum_{i \in \mathcal{D}_n} \sigma_i (P_u^N(i) \right. \\
 & \left. \left. - p_\infty(F_u^N(i)))^2 \right] du \right) \leq \mathbb{E} [V_{\hat{\mathcal{D}}} (H_0^N, P_0^N, F_0^N)] + tb^N c_0.
 \end{aligned} \tag{4.94}$$

The set $\hat{\mathcal{D}} \subseteq \mathcal{D}$ was arbitrarily chosen and thus, this completes the proof of Theorem 4.1.9. \square

4.1.4.4 Proof of Theorem 3.1.2

In this section, the proof of Theorem 3.1.2 will be given, after providing some useful lemmas. For convenience of the reader, Lemma 3.3 of Klenke and Mytnik (2012) is restated.

Lemma 4.1.10. *Let \mathcal{D} be a countable set, let $\sigma \in (0, \infty)^\mathcal{D}$ such that $\sum_{i \in \mathcal{D}} \sigma_i < \infty$, and let $l_\sigma^1 := \{z \in \mathbb{R}^\mathcal{D} : \|z\|_\sigma := \sum_{i \in \mathcal{D}} \sigma_i z_i < \infty\}$. A subset $K \subseteq l_\sigma^1$ is relatively compact if and only if*

$$(i) \sup_{x \in K} \|x\|_\sigma < \infty$$

$$(ii) \text{ for every } \varepsilon \in (0, \infty) \text{ there exists a finite subset } \mathcal{E} \subseteq \mathcal{D} \text{ such that } \sup_{x \in K} \|x \mathbf{1}_{\mathcal{D} \setminus \mathcal{E}}\|_\sigma < \varepsilon.$$

Lemma 4.1.11. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let \mathcal{D} be a countable set, let $\sigma \in (0, \infty)^\mathcal{D}$ such that $\sum_{i \in \mathcal{D}} \sigma_i < \infty$, let $l_\sigma^1 := \{z \in \mathbb{R}^\mathcal{D} : \|z\|_\sigma := \sum_{i \in \mathcal{D}} \sigma_i z_i < \infty\}$, let $E_2 := l_\sigma^1 \cap [0, \infty)^\mathcal{D}$, let I be a set, and let $Z^i : \Omega \rightarrow E_2$, $i \in I$, be a family of random variables. Assume that $\sup_{i \in I} \mathbb{E}[\|Z^i\|_\sigma] < \infty$ and $\inf_{S \subseteq \mathcal{D}, |S| < \infty} \sup_{i \in I} \sum_{k \in \mathcal{D} \setminus S} \sigma_k \mathbb{E}[Z_k^i] = 0$. Then the family $\{Z^i : i \in I\}$ is relatively compact in E_2 .*

Proof. Fix $\varepsilon \in (0, \infty)$. For each $m \in \mathbb{N}$ by assumption there exists a set $\mathcal{S}_{m, \varepsilon} \subseteq \mathcal{D}$ such that

$$\sup_{i \in I} \sum_{k \in \mathcal{D} \setminus \mathcal{S}_{m, \varepsilon}} \sigma_k \mathbb{E}[Z_k^i] < \frac{\varepsilon}{2m^2(m+1)}. \tag{4.95}$$

Define the set $K_\varepsilon \subseteq E_2$ by

$$K_\varepsilon := \left\{ x \in E_2 : \|x\|_\sigma \leq \frac{2 \sup_{i \in I} \mathbb{E}[\|Z^i\|_\sigma]}{\varepsilon}, \sup_{m \in \mathbb{N}} \left\{ m \sum_{k \in \mathcal{D} \setminus \mathcal{S}_{m,\varepsilon}} \sigma_k |x_k| \right\} \leq 1 \right\}. \quad (4.96)$$

Due to the Heine-Borel theorem, Lemma 4.1.10 can be applied and yields relative compactness of K_ε . By Markov's inequality it now holds that

$$\begin{aligned} \sup_{i \in I} \mathbb{P}[Z^i \notin \overline{K_\varepsilon}] &\leq \sup_{i \in I} \mathbb{P}[Z^i \notin K_\varepsilon] \\ &\leq \sup_{i \in I} \mathbb{P}\left[\|Z^i\|_\sigma > \frac{2 \sup_{j \in I} \mathbb{E}[\|Z^j\|_\sigma]}{\varepsilon}\right] + \sup_{i \in I} \sum_{m=1}^{\infty} \mathbb{P}\left[\sum_{k \in \mathcal{D} \setminus \mathcal{S}_{m,\varepsilon}} \sigma_k Z_k^i > \frac{1}{m}\right] \\ &\leq \frac{\varepsilon \sup_{i \in I} \mathbb{E}[\|Z^i\|_\sigma]}{2 \sup_{j \in I} \mathbb{E}[\|Z^j\|_\sigma]} + \sum_{m=1}^{\infty} m \sup_{i \in I} \sum_{k \in \mathcal{D} \setminus \mathcal{S}_{m,\varepsilon}} \sigma_k \mathbb{E}[Z_k^i] \leq \frac{\varepsilon}{2} + \sum_{m=1}^{\infty} m \frac{\varepsilon}{2m^2(m+1)} = \varepsilon. \end{aligned} \quad (4.97)$$

Since ε was arbitrarily chosen it follows that $\{Z^i : i \in I\}$ is tight in E_2 . Due to Prohorov's theorem (e.g., Ethier and Kurtz, 1986, Theorem 3.2.2) the claim follows. \square

Lemma 4.1.12. *Assume the setting of Section 4.1.4.1 and assume for all $N \in \mathbb{N}$ that $\sum_{i \in \mathcal{D}} \sigma_i \mathbb{E}[H_0^N(i)] < \infty$. For all $n \in \mathbb{N}$ denote by m^n the n -fold matrix product of m . Then it holds for all $t \in [0, \infty)$, all $i \in \mathcal{D}$, and all $N \in \mathbb{N}$ that*

$$\mathbb{E}[H_t^N(i)] \leq \mathbb{E}\left[\sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-t\kappa_H^N} \frac{(t\kappa_H^N)^n}{n!} m^n(i, j) H_0^N(j)\right] + \frac{K}{2} \left(1 + \sqrt{1 + \frac{4\bar{L}_H}{K\lambda}}\right). \quad (4.98)$$

Proof. For every $n \in \mathbb{N}$ and every $i, j \in \mathcal{D}$ it holds that $m^n(i, j) \in [0, 1]$. Hence, for all $T \in [0, \infty)$ and all $i, j \in \mathcal{D}$ it follows that

$$\sum_{n=0}^{\infty} \sup_{t \in [0, T]} e^{-t\kappa_H^N} \frac{t^n}{n!} m^n(i, j) < \infty. \quad (4.99)$$

Now, define for all $t \in [0, \infty)$ and all $i, j \in \mathcal{D}$

$$m_t(i, j) := \sum_{n=0}^{\infty} e^{-t\kappa_H^N} \frac{t^n}{n!} m^n(i, j). \quad (4.100)$$

By (4.99) and using dominated convergence, it follows for all $t \in [0, \infty)$ and all $i, j \in \mathcal{D}$

that

$$\begin{aligned}
 \frac{d}{dt} m_t(i, j) &= -m_t(i, j) + \sum_{n=1}^{\infty} e^{-t} \frac{t^{n-1}}{(n-1)!} m^n(i, j) \\
 &= -m_t(i, j) + \sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} m^{n+1}(i, j) \\
 &= -m_t(i, j) + \sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} \sum_{k \in \mathcal{D}} m^n(i, k) m(k, j) = \sum_{k \in \mathcal{D}} m_t(i, k) (m(k, j) - \mathbb{1}_{j=k}).
 \end{aligned} \tag{4.101}$$

Furthermore, note that for all $t \in [0, \infty)$ and all $i \in \mathcal{D}$ it holds that

$$\sum_{j \in \mathcal{D}} m_t(i, j) = \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} m^n(i, j) = \sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} = 1. \tag{4.102}$$

For all $t \in [0, \infty)$, $s \in [0, t]$, $i \in \mathcal{D}$, $N \in \mathbb{N}$ define

$$Y_s^{N,t}(i) := \sum_{j \in \mathcal{D}} m_{(t-s)\kappa_H^N}(i, j) H_s^N(j). \tag{4.103}$$

Observe that since for all $i, j \in \mathcal{D}$ it holds that $m_0(i, j) = \mathbb{1}_{i=j}$ it follows for all $t \in [0, \infty)$, all $i \in \mathcal{D}$, and all $N \in \mathbb{N}$ that

$$Y_t^{N,t}(i) = H_t^N(i). \tag{4.104}$$

Furthermore, using (4.10) yields for all $t \in [0, \infty)$ and all $N \in \mathbb{N}$ that

$$\begin{aligned}
 \sum_{i \in \mathcal{D}} \sigma_i \mathbb{E} [Y_0^{N,t}(i)] &= \sum_{i \in \mathcal{D}} \sigma_i \mathbb{E} \left[\sum_{j \in \mathcal{D}} m_{t\kappa_H^N}(i, j) H_0^N(j) \right] \\
 &= \sum_{i \in \mathcal{D}} \sigma_i \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-t\kappa_H^N} \frac{(t\kappa_H^N)^n}{n!} m^n(i, j) \mathbb{E} [H_0^N(j)] \\
 &= \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-t\kappa_H^N} \frac{(t\kappa_H^N)^n}{n!} \mathbb{E} [H_0^N(j)] \sum_{i \in \mathcal{D}} \sigma_i m^n(i, j) \\
 &\leq \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-t\kappa_H^N} \frac{(t\kappa_H^N)^n}{n!} \mathbb{E} [H_0^N(j)] c^n \sigma_j = \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-t\kappa_H^N} \frac{(t\kappa_H^N c)^n}{n!} \mathbb{E} [H_0^N(j)] \sigma_j \\
 &= e^{t\kappa_H^N(c-1)} \mathbb{E} [\|H_0^N\|_{\sigma}].
 \end{aligned} \tag{4.105}$$

For all $t \in [0, \infty)$, $s \in [0, t]$, $i, j \in \mathcal{D}$, $N \in \mathbb{N}$ it follows from (4.101) that

$$\frac{d}{ds} m_{(t-s)\kappa_H^N}(i, j) = -\kappa_H^N \sum_{k \in \mathcal{D}} m_{(t-s)\kappa_H^N}(i, k) (m(k, j) - \mathbb{1}_{j=k}). \quad (4.106)$$

For $t \in [0, \infty)$, $N, l \in \mathbb{N}$, $i \in \mathcal{D}$ define

$$\tau_l^{N,t}(i) := \inf \left(\{u \in [0, t] : Y_u^{N,t}(i) > l\} \cup \infty \right). \quad (4.107)$$

Using the fact that for all $t \in [0, \infty)$, all $u \in [0, t]$, all $N \in \mathbb{N}$, and all $i, j \in \mathcal{D}$ it holds that $m_{(t-u)\kappa_H^N}(i, j) \in [0, 1]$ implies for all $t \in [0, \infty)$, all $s \in [0, t]$, all $N, l \in \mathbb{N}$, and all $i \in \mathcal{D}$ that

$$\begin{aligned} & \int_0^{s \wedge \tau_l^{N,t}(i)} \sum_{j \in \mathcal{D}} \left(m_{(t-u)\kappa_H^N}(i, j) \sqrt{\beta_H^N H_u^N(j)} \right)^2 du \\ & \leq \int_0^{s \wedge \tau_l^{N,t}(i)} \sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) \beta_H^N H_u^N(j) du = \int_0^{s \wedge \tau_l^{N,t}(i)} \beta_H^N Y_u^{N,t}(i) du \\ & \leq \int_0^s \beta_H^N Y_{u \wedge \tau_l^{N,t}(i)}^{N,t}(i) du \leq t \beta_H^N l. \end{aligned} \quad (4.108)$$

For all $t \in [0, \infty)$, $s \in [0, t]$, $i \in \mathcal{D}$, $N \in \mathbb{N}$ using Itô's lemma with (4.102) and (4.106)

yields \mathbb{P} -a.s.

$$\begin{aligned}
 Y_s^{N,t}(i) - Y_0^{N,t}(i) &= \int_0^s \sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) \left(\kappa_H^N \sum_{k \in \mathcal{D}} m(j, k) H_u^N(k) + (\lambda - \kappa_H^N \right. \\
 &\quad \left. - \alpha^N F_t^N(j)) H_u^N(j) - \frac{\lambda}{K} (H_u^N(j))^2 - \delta H_u^N(j) P_u^N(j) + \iota_H^N \right) \\
 &\quad - \sum_{j \in \mathcal{D}} \kappa_H^N \sum_{k \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, k) (m(k, j) - \mathbb{1}_{j=k}) H_u^N(j) du \\
 &\quad + \sum_{j \in \mathcal{D}} \int_0^s m_{(t-u)\kappa_H^N}(i, j) \sqrt{\beta_H^N H_u^N(j)} dW_u^{N,H}(j) \\
 &= \int_0^s \sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) \left((\lambda - \alpha^N F_t^N(j)) H_u^N(j) - \frac{\lambda}{K} (H_u^N(j))^2 \right. \\
 &\quad \left. - \delta H_u^N(j) P_u^N(j) \right) + \iota_H^N du + \sum_{j \in \mathcal{D}} \int_0^s m_{(t-u)\kappa_H^N}(i, j) \sqrt{\beta_H^N H_u^N(j)} dW_u^{N,H}(j) \\
 &\leq \int_0^s \sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) \left(\lambda H_u^N(j) - \frac{\lambda}{K} (H_u^N(j))^2 \right) + \iota_H^N du \\
 &\quad + \sum_{j \in \mathcal{D}} \int_0^s m_{(t-u)\kappa_H^N}(i, j) \sqrt{\beta_H^N H_u^N(j)} dW_u^{N,H}(j).
 \end{aligned} \tag{4.109}$$

Thus, using (4.108) and (4.109) it follows for all $t \in [0, \infty)$, all $s \in [0, t]$, all $i \in \mathcal{D}$, and all $N, l \in \mathbb{N}$ that

$$\begin{aligned}
 &\mathbb{E} \left[Y_{s \wedge \tau_l^{N,t}(i)}^{N,t}(i) \right] - \mathbb{E} \left[Y_0^{N,t}(i) \right] \\
 &\leq \mathbb{E} \left[\int_0^{s \wedge \tau_l^{N,t}(i)} \sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) \lambda H_u^N(j) + \iota_H^N du \right] \\
 &\leq \mathbb{E} \left[\int_0^s \sum_{j \in \mathcal{D}} m_{(t-u \wedge \tau_l^{N,t}(i))\kappa_H^N}(i, j) \lambda H_{u \wedge \tau_l^{N,t}(i)}^N(j) + \iota_H^N du \right] \\
 &= \int_0^s \lambda \mathbb{E} \left[Y_{u \wedge \tau_l^{N,t}(i)}^{N,t}(i) \right] + \iota_H^N du \leq t \iota_H^N + \lambda \int_0^s \mathbb{E} \left[Y_{u \wedge \tau_l^{N,t}(i)}^{N,t}(i) \right] du.
 \end{aligned} \tag{4.110}$$

Now, using Gronwall's lemma (e.g., Klenke, 2008), it follows for all $t \in [0, \infty)$, all $s \in [0, t]$, all $i \in \mathcal{D}$, and all $N, l \in \mathbb{N}$ that

$$\mathbb{E} \left[Y_{s \wedge \tau_l^{N,t}(i)}^{N,t}(i) \right] \leq \left(\mathbb{E} \left[Y_0^{N,t}(i) \right] + t \iota_H^N \right) e^{\lambda s} \leq \left(\mathbb{E} \left[Y_0^{N,t}(i) \right] + t \iota_H^N \right) e^{\lambda t}. \tag{4.111}$$

For all $t \in [0, \infty)$, $N \in \mathbb{N}$, $i \in \mathcal{D}$ the \mathbb{P} -a.s. continuous paths of $(Y_u^{N,t}(i))_{u \in [0,t]}$ imply $\mathbb{P}\left[\sup_{u \in [0,t]} Y_u^{N,t}(i) < \infty\right] = 1$. Hence, it holds for all $t \in [0, \infty)$, all $N \in \mathbb{N}$, and all $i \in \mathcal{D}$ that

$$\mathbb{P}\left[\lim_{l \rightarrow \infty} \tau_l^{N,t}(i) = \infty\right] = 1. \quad (4.112)$$

Using the assumption that for all $N \in \mathbb{N}$ it holds that $\sum_{i \in \mathcal{D}} \sigma_i \mathbb{E}[H_0^N(i)] < \infty$ together with (4.104), (4.105), (4.111), and (4.112) with Fatou's lemma implies for all $t \in [0, \infty)$ and all $N \in \mathbb{N}$ that

$$\begin{aligned} \sum_{i \in \mathcal{D}} \sigma_i \mathbb{E}[H_t^N(i)] &= \sum_{i \in \mathcal{D}} \sigma_i \mathbb{E}[Y_t^{N,t}(i)] = \sum_{i \in \mathcal{D}} \sigma_i \mathbb{E}\left[\lim_{l \rightarrow \infty} Y_{t \wedge \tau_l^{N,t}(i)}^{N,t}(i)\right] \\ &\leq \sum_{i \in \mathcal{D}} \sigma_i \liminf_{l \rightarrow \infty} \mathbb{E}\left[Y_{t \wedge \tau_l^{N,t}(i)}^{N,t}(i)\right] \leq \sum_{i \in \mathcal{D}} \sigma_i \liminf_{l \rightarrow \infty} \left(\mathbb{E}[Y_0^{N,t}(i)] + t\ell_H^N\right) e^{\lambda t} \\ &= \sum_{i \in \mathcal{D}} \sigma_i \left(\mathbb{E}[Y_0^{N,t}(i)] + t\ell_H^N\right) e^{\lambda t} \leq \left(e^{t\kappa_H^N(c-1)} \mathbb{E}[\|H_0^N\|_\sigma] + \sum_{i \in \mathcal{D}} \sigma_i t\ell_H^N\right) e^{\lambda t} \\ &< \infty. \end{aligned} \quad (4.113)$$

Using the fact that for all $t \in [0, \infty)$, all $u \in [0, t]$, all $N \in \mathbb{N}$, and all $i, j \in \mathcal{D}$ it holds that $m_{(t-u)\kappa_H^N}(i, j) \in [0, 1]$ this implies for all $t \in [0, \infty)$, all $s \in [0, t]$, all $N \in \mathbb{N}$, and all $i \in \mathcal{D}$, that

$$\begin{aligned} &\mathbb{E}\left[\sum_{j \in \mathcal{D}} \int_0^s \left(m_{(t-u)\kappa_H^N}(i, j) \sqrt{\beta_H^N H_u^N(j)}\right)^2 du\right] \\ &= \int_0^s \mathbb{E}\left[\sum_{j \in \mathcal{D}} \left(m_{(t-u)\kappa_H^N}(i, j) \sqrt{\beta_H^N H_u^N(j)}\right)^2\right] du \\ &\leq \beta_H^N \int_0^s \mathbb{E}\left[\sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) H_u^N(j)\right] du = \beta_H^N \int_0^s \mathbb{E}[Y_u^{N,t}(i)] du < \infty. \end{aligned} \quad (4.114)$$

Thus, taking expectations in (4.109) gives for all $t \in [0, \infty)$, all $s \in [0, t]$, all $i \in \mathcal{D}$, and all

$N \in \mathbb{N}$ using Jensen's inequality

$$\begin{aligned}
 & \mathbb{E} [Y_s^{N,t}(i)] - \mathbb{E} [Y_0^{N,t}(i)] \\
 & \leq \int_0^s \left(\lambda \mathbb{E} [Y_u^{N,t}(i)] - \frac{\lambda}{K} \mathbb{E} \left[\sum_{j \in \mathcal{D}} m_{(t-u)\kappa_H^N}(i, j) (H_u^N(j))^2 \right] \right) + \iota_H^N du \\
 & \leq \int_0^s \left(\lambda \mathbb{E} [Y_u^{N,t}(i)] - \frac{\lambda}{K} \mathbb{E} [(Y_u^{N,t}(i))^2] \right) + \iota_H^N du \\
 & \leq \int_0^s \left(\lambda \mathbb{E} [Y_u^{N,t}(i)] - \frac{\lambda}{K} (\mathbb{E} [Y_u^{N,t}(i)])^2 \right) + \bar{\iota}_H du.
 \end{aligned} \tag{4.115}$$

For $t \in [0, \infty)$, $i \in \mathcal{D}$, $N \in \mathbb{N}$ let $z^{N,t}(i) : [0, \infty) \rightarrow \mathbb{R}$ be a process that for all $s \in [0, \infty)$ satisfies

$$z_s^{N,t}(i) = z_0^{N,t}(i) + \int_0^s \left(\lambda z_u^{N,t}(i) - \frac{\lambda}{K} (z_u^{N,t}(i))^2 + \bar{\iota}_H \right) du \tag{4.116}$$

with $z_0^{N,t}(i) = \mathbb{E} [Y_0^{N,t}(i)]$ where uniqueness follows from local Lipschitz continuity. Define $c_1 := \frac{K}{2} + \sqrt{\frac{K^2}{4} + \frac{K\bar{\iota}_H}{\lambda}} \in (0, \infty)$. Using classical comparison results from the theory of ODEs, the above computation shows for all $N \in \mathbb{N}$, all $i \in \mathcal{D}$, and all $t \in [0, \infty)$ that

$$\begin{aligned}
 \mathbb{E} [H_t^N(i)] &= \mathbb{E} [Y_t^{N,t}(i)] \leq z_t^{N,t}(i) \leq \max \left\{ \mathbb{E} [Y_0^{N,t}(i)], \limsup_{s \rightarrow \infty} z_s^{N,t}(i) \right\} \\
 &= \max \left\{ \mathbb{E} [Y_0^{N,t}(i)], c_1 \right\} \leq \mathbb{E} \left[\sum_{j \in \mathcal{D}} m_{t\kappa_H^N}(i, j) H_0^N(j) \right] + c_1.
 \end{aligned} \tag{4.117}$$

This completes the proof of Lemma 4.1.12. \square

Proof of Theorem 3.1.2. The result will be proved using stochastic averaging (see Kurtz, 1992, Theorem 2.1). First the conditions of the aforementioned theorem are verified. Note that $E_1 = [0, 1]^\mathcal{D}$ and $E_2 = l_\sigma^1 \cap [0, \infty)^\mathcal{D}$ are complete separable metric spaces. Tychonoff's theorem implies that E_1 is compact. Since for all $N \in \mathbb{N}$ and all $t \in [0, \infty)$ the random variable F_{tN}^N takes values in the compact space E_1 , the compact containment condition holds for $\{ (F_{tN}^N)_{t \in [0, \infty)} : N \in \mathbb{N} \}$. Now Lemma 4.1.11 is used to show for each $T \in [0, \infty)$ that the family $\{ H_{tN}^N : t \in [0, T], N \in \mathbb{N} \}$ is relatively compact in E_2 . From Lemma 4.1.6 and the assumption $\sup_{N \in \mathbb{N}} \mathbb{E} \left[\left\| (H_0^N + P_0^N)^4 \right\| \right] < \infty$ it follows that

$$\sup_{N \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} [\|H_{tN}^N\|_\sigma] < \infty. \tag{4.118}$$

Define $\mathcal{D}_0 := \emptyset$ and for all $n \in \mathbb{N}$ let $\mathcal{D}_n \subseteq \mathcal{D}$ be a set with $|\mathcal{D}_n| = \min\{n, |\mathcal{D}|\}$ and $\mathcal{D}_n \supseteq \mathcal{D}_{n-1}$. Define $c_1 := \frac{K}{2} \left(1 + \sqrt{1 + \frac{4\bar{t}_H}{K\lambda}}\right)$. From Lemma 4.1.12 with the assumption that $\sum_{i \in \mathcal{D}} \sup_{N \in \mathbb{N}} \sigma_i \mathbb{E} [H_0^N(i)] < \infty$ it follows for all $T \in [0, \infty)$ that

$$\begin{aligned}
 & \sum_{i \in \mathcal{D}} \sigma_i \sup_{N \in \mathbb{N}} \sup_{t \in [0, T]} \mathbb{E} [H_{tN}^N(i)] \\
 & \leq \sum_{i \in \mathcal{D}} \sigma_i \sup_{N \in \mathbb{N}} \sup_{t \in [0, T]} \left(\sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} e^{-tN\kappa_H^N} \frac{(tN\kappa_H^N)^n}{n!} m^n(i, j) \mathbb{E} [H_0^N(j)] + c_1 \right) \\
 & \leq \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} \left(\sum_{i \in \mathcal{D}} \sigma_i m^n(i, j) \right) \sup_{N \in \mathbb{N}} \sup_{t \in [0, TN\kappa_H^N]} e^{-t\frac{tn}{n!}} \mathbb{E} [H_0^N(j)] + c_1 \sum_{i \in \mathcal{D}} \sigma_i \quad (4.119) \\
 & \leq \sum_{j \in \mathcal{D}} \sum_{n=0}^{\infty} c^n \sigma_j \sup_{N \in \mathbb{N}} \frac{(TN\kappa_H^N)^n}{n!} \mathbb{E} [H_0^N(j)] + c_1 \|\mathbf{1}\|_{\sigma} \\
 & \leq e^{cT} \sup_{M \in \mathbb{N}} M \kappa_H^M \sum_{j \in \mathcal{D}} \sigma_j \sup_{N \in \mathbb{N}} \mathbb{E} [H_0^N(j)] + c_1 \|\mathbf{1}\|_{\sigma} < \infty.
 \end{aligned}$$

Now the dominated convergence theorem yields for all $T \in [0, \infty)$ that

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} \sup_{N \in \mathbb{N}} \sup_{t \in [0, T]} \sum_{k \in \mathcal{D} \setminus \mathcal{D}_n} \sigma_k \mathbb{E} [H_{tN}^N(k)] \leq \lim_{n \rightarrow \infty} \sum_{k \in \mathcal{D} \setminus \mathcal{D}_n} \sup_{N \in \mathbb{N}} \sup_{t \in [0, T]} \sigma_k \mathbb{E} [H_{tN}^N(k)] \\
 & = 0.
 \end{aligned} \quad (4.120)$$

Hence, for all $T \in [0, \infty)$ Lemma 4.1.11 can be applied to the family $\{H_{tN}^N : t \in [0, T], N \in \mathbb{N}\}$ to show that it is relatively compact in E_2 . Denote by $C_b(E_1, \mathbb{R})$ the set of bounded, continuous real-valued functions on E_1 and by $C_b^2(E_1, \mathbb{R})$ the set of all real-valued functions on E_1 that are twice continuously differentiable and bounded, with bounded first and second order partial derivatives. For $f \in C_b^2(E_1, \mathbb{R})$ let $c_f \in (0, \infty)$ be such that for all $x \in E_1$ and all $i \in \mathcal{D}$ it holds that $\left| \frac{df}{dx_i}(x) \right| + \left| \frac{d^2f}{dx_i^2}(x) \right| \leq c_f$. Define

$$\text{Dom}(\mathcal{A}) := \{f \in C_b^2(E_1, \mathbb{R}) : f \text{ depends only on finitely many coordinates}\} \quad (4.121)$$

and for any $f \in \text{Dom}(\mathcal{A})$ denote by \mathcal{D}_f the finite set of coordinates that f depends on. Due to the Stone-Weierstrass theorem (e.g., Klenke, 2008, Theorem 15.2) it follows that $\text{Dom}(\mathcal{A})$ is dense in $C_b(E_1, \mathbb{R})$ in the topology of uniform convergence. Denote by $C(E_1 \times E_2, \mathbb{R})$ the set of real-valued continuous functions on $E_1 \times E_2$ and define the operator $\mathcal{A}_1 : \text{Dom}(\mathcal{A}) \rightarrow C(E_1 \times E_2, \mathbb{R})$ for all $f \in \text{Dom}(\mathcal{A})$, all $x \in E_1$, and all $y \in E_2$ by $(\mathcal{A}_1 f)(x, y) :=$

$\sum_{i \in \mathcal{D}} \mathbb{1}_{y_i > 0} \left(\left[\kappa_H \sum_{j \in \mathcal{D}} \left(m(i, j) \frac{y_j}{y_i} (x_j - x_i) \right) - \alpha x_i (1 - x_i) \right] \frac{df}{dx_i}(x) + \frac{1}{2} \beta_H \frac{x_i(1-x_i)}{y_i} \frac{d^2 f}{dx_i^2}(x) \right)$. For all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all $t \in [0, \infty)$ define

$$\begin{aligned}
 \varepsilon_f^N(t) &:= \frac{1}{N} \int_0^t (\mathcal{A}_1 f)(F_u^N, H_u^N) du - \sum_{i \in \mathcal{D}} \int_0^t \frac{df}{dx_i}(F_u^N) \left[\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (F_u^N(j) \right. \\
 &\quad \left. - F_u^N(i)) \frac{H_u^N(j)}{H_u^N(i)} - \alpha^N F_u^N(i) (1 - F_u^N(i)) \right] + \frac{1}{2} \frac{d^2 f}{dx_i^2}(F_u^N) \frac{\beta_H F_u^N(i)(1-F_u^N(i))}{N H_u^N(i)} du.
 \end{aligned} \tag{4.122}$$

From Itô's lemma and Lemma 4.1.5 it follows for all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all $t \in [0, \infty)$ that \mathbb{P} -a.s.

$$\begin{aligned}
 f(F_t^N) - f(F_0^N) &= \sum_{i \in \mathcal{D}} \int_0^t \frac{df}{dx_i}(F_u^N) dF_u^N(i) \\
 &\quad + \frac{1}{2} \sum_{i, j \in \mathcal{D}} \int_0^t \left(\frac{d^2 f}{dx_i dx_j}(F_u^N) \right) d \langle F^N(i), F^N(j) \rangle_u \\
 &= \sum_{i \in \mathcal{D}} \int_0^t \frac{df}{dx_i}(F_u^N) \left[\kappa_H^N \sum_{j \in \mathcal{D}} m(i, j) (F_u^N(j) - F_u^N(i)) \frac{H_u^N(j)}{H_u^N(i)} \right. \\
 &\quad \left. - \alpha^N F_u^N(i) (1 - F_u^N(i)) \right] + \frac{1}{2} \frac{d^2 f}{dx_i^2}(F_u^N) \frac{\beta_H F_u^N(i)(1-F_u^N(i))}{H_u^N(i)} du \\
 &\quad + \sum_{i \in \mathcal{D}} \int_0^t \frac{df}{dx_i}(F_u^N) \sqrt{\frac{\beta_H F_u^N(i)(1-F_u^N(i))}{H_u^N(i)}} dW_u^{F, N}(i).
 \end{aligned} \tag{4.123}$$

Hence, it follows for all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all $t \in [0, \infty)$ that \mathbb{P} -a.s.

$$\begin{aligned}
 f(F_{tN}^N) - \int_0^t (\mathcal{A}_1 f)(F_{uN}^N, H_{uN}^N) du + \varepsilon_f^N(tN) \\
 = f(F_0^N) + \sum_{i \in \mathcal{D}} \int_0^t \frac{df}{dx_i}(F_{uN}^N) \sqrt{\frac{\beta_H F_{uN}^N(i)(1-F_{uN}^N(i))}{H_{uN}^N(i)}} dW_{uN}^{F, N}(i).
 \end{aligned} \tag{4.124}$$

From Tonelli's theorem and Lemma 4.1.7 it follows for all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all

$t \in [0, \infty)$ that

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^t \left(\sum_{i \in \mathcal{D}} \frac{df}{dx_i}(F_u^N) \sqrt{\frac{\beta_H^N F_u^N(i)(1-F_u^N(i))}{H_u^N(i)}} \right)^2 du \right] \\
 & \leq t |\mathcal{D}_f| c_f^2 \bar{\beta}_H \max_{i \in \mathcal{D}_f} \sup_{M \in \mathbb{N}} \sup_{u \in [0, \infty)} \mathbb{E} \left[\frac{1}{H_u^M(i)} \right] \\
 & \leq t |\mathcal{D}_f| c_f^2 \bar{\beta}_H \max_{i \in \mathcal{D}_f} \frac{1}{\sigma_i} \sup_{M \in \mathbb{N}} \sup_{u \in [0, \infty)} \mathbb{E} \left[\left\| \frac{1}{H_u^M} \right\|_\sigma \right] < \infty.
 \end{aligned} \tag{4.125}$$

Thus for all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all $t \in [0, \infty)$ the left-hand side of (4.124) is a martingale. Next, for all $f \in \text{Dom}(\mathcal{A})$ and all $T \in [0, \infty)$ it holds that

$$\begin{aligned}
 & \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T |(\mathcal{A}_1 f)(F_{tN}^N, H_{tN}^N)|^{\frac{4}{3}} dt \right] \\
 & = \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T \left| \sum_{i \in \mathcal{D}_f} \left(\kappa_H \sum_{j \in \mathcal{D}} \left(m(i, j) \frac{H_{tN}^N(j)}{H_{tN}^N(i)} (F_{tN}^N(j) - F_{tN}^N(i)) \right) \right. \right. \right. \\
 & \quad \left. \left. - \alpha F_{tN}^N(i) (1 - F_{tN}^N(i)) \right) \frac{df}{dx_i}(F_{tN}^N) + \frac{1}{2} \sum_{i \in \mathcal{D}_f} \beta_H \frac{F_{tN}^N(i)(1-F_{tN}^N(i))}{H_{tN}^N(i)} \frac{d^2 f}{dx_i^2}(F_{tN}^N) \right|^{\frac{4}{3}} dt \right] \\
 & \leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T \left(\sum_{i \in \mathcal{D}_f} \left(\left| \kappa_H \sum_{j \in \mathcal{D}} \left(m(i, j) \frac{H_{tN}^N(j)}{H_{tN}^N(i)} \right) c_f \right| + |\alpha c_f| \right. \right. \right. \\
 & \quad \left. \left. + \left| \frac{1}{2} \beta_H \frac{1}{H_{tN}^N(i)} c_f \right| \right) \right)^{\frac{4}{3}} dt \right].
 \end{aligned} \tag{4.126}$$

Young's inequality and Jensen's inequality imply for all $f \in \text{Dom}(\mathcal{A})$ and all $T \in [0, \infty)$ that

$$\begin{aligned}
 & \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T |(\mathcal{A}_1 f)(F_{tN}^N, H_{tN}^N)|^{\frac{4}{3}} dt \right] \leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T \left(\sum_{i \in \mathcal{D}_f} \left(\frac{2}{3} \left(\kappa_H c_f \frac{1}{H_{tN}^N(i)} \right)^{\frac{3}{2}} \right. \right. \right. \\
 & \quad \left. \left. + \frac{1}{3} \left(\sum_{j \in \mathcal{D}} m(i, j) H_{tN}^N(j) \right)^3 + \alpha c_f + \frac{1}{2} \beta_H \frac{1}{H_{tN}^N(i)} c_f \right) \right)^{\frac{4}{3}} dt \right] \\
 & \leq \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T \frac{(4|\mathcal{D}_f|)^{\frac{1}{3}}}{\min_{k \in \mathcal{D}_f} \{\sigma_k\}} \sum_{i \in \mathcal{D}_f} \sigma_i \left(\left(\frac{2}{3} \right)^{\frac{4}{3}} \left(\kappa_H c_f \frac{1}{H_{tN}^N(i)} \right)^2 \right. \right. \\
 & \quad \left. \left. + \left(\frac{1}{3} \right)^{\frac{4}{3}} \left(\sum_{j \in \mathcal{D}} m(i, j) H_{tN}^N(j) \right)^4 + (\alpha c_f)^{\frac{4}{3}} + \left(\frac{1}{2} \beta_H \frac{1}{H_{tN}^N(i)} c_f \right)^{\frac{4}{3}} \right) dt \right].
 \end{aligned} \tag{4.127}$$

Lemma 4.1.5, Tonelli's theorem, and Lemmas 4.1.6 and 4.1.7 imply for all $f \in \text{Dom}(\mathcal{A})$ and all $T \in [0, \infty)$ that

$$\begin{aligned} & \sup_{N \in \mathbb{N}} \mathbb{E} \left[\int_0^T |(\mathcal{A}_1 f)(F_{tN}^N, H_{tN}^N)|^{\frac{4}{3}} dt \right] \\ & \leq \sup_{N \in \mathbb{N}} \frac{(4|\mathcal{D}_f|)^{\frac{2}{3}}}{\min_{k \in \mathcal{D}_f} \{\sigma_k\}} \int_0^T \left(\frac{4}{9}\right)^{\frac{2}{3}} (\kappa_H c_f)^2 \mathbb{E} \left[\left\| \frac{1}{(H_{tN}^N)^2} \right\|_{\sigma} \right] + \left(\frac{1}{9}\right)^{\frac{2}{3}} c \mathbb{E} \left[\left\| (H_{tN}^N)^4 \right\|_{\sigma} \right. \\ & \quad \left. + (\alpha c_f)^{\frac{4}{3}} \|\mathbb{1}\|_{\sigma} \right] + \left(\frac{1}{4}\right)^{\frac{2}{3}} (\beta_H c_f)^{\frac{4}{3}} \mathbb{E} \left[\left\| \left(\frac{1}{H_{tN}^N}\right)^{\frac{4}{3}} \right\|_{\sigma} \right] dt < \infty. \end{aligned} \quad (4.128)$$

Furthermore, for all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all $T \in [0, \infty)$ it holds that

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \in [0, T]} |\varepsilon_f^N(tN)| \right] = \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i \in \mathcal{D}_f} \int_0^t \frac{df}{dx_i}(F_{uN}^N) \left[(\kappa_H \right. \right. \right. \\ & \quad \left. \left. - N\kappa_H^N) \sum_{j \in \mathcal{D}} m(i, j) (F_{uN}^N(j) - F_{uN}^N(i)) \frac{H_{uN}^N(j)}{H_{uN}^N(i)} + (\alpha - N\alpha^N) F_{uN}^N(i) (1 \right. \right. \\ & \quad \left. \left. - F_{uN}^N(i)) \right] + \frac{1}{2} \frac{d^2 f}{dx_i^2}(F_{uN}^N) (\beta_H - N\beta_H^N) \frac{F_{uN}^N(i)(1 - F_{uN}^N(i))}{H_{uN}^N(i)} du \right] \\ & \leq \mathbb{E} \left[\int_0^T \sum_{i \in \mathcal{D}_f} c_f \left(|\kappa_H - N\kappa_H^N| \sum_{j \in \mathcal{D}} m(i, j) \frac{H_{uN}^N(j)}{H_{uN}^N(i)} + |\alpha - N\alpha^N| \right. \right. \\ & \quad \left. \left. + \frac{1}{2} |\beta_H - N\beta_H^N| \frac{1}{H_{uN}^N(i)} \right) du \right]. \end{aligned} \quad (4.129)$$

Young's inequality, Lemma 4.1.5, and Tonelli's theorem imply for all $f \in \text{Dom}(\mathcal{A})$, all $N \in \mathbb{N}$, and all $T \in [0, \infty)$ that

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \in [0, T]} |\varepsilon_f^N(tN)| \right] \leq \mathbb{E} \left[\int_0^T \sum_{i \in \mathcal{D}_f} \frac{\sigma_i c_f}{\min_{k \in \mathcal{D}_f} \{\sigma_k\}} \left(\frac{|\kappa_H - N\kappa_H^N|}{2} \left(\left(\frac{1}{H_{uN}^N(i)} \right)^2 \right. \right. \right. \\ & \quad \left. \left. + \left(\sum_{j \in \mathcal{D}} m(i, j) H_{uN}^N(j) \right)^2 \right) + |\alpha - N\alpha^N| + \frac{|\beta_H - N\beta_H^N|}{2} \frac{1}{H_{uN}^N(i)} \right) du \right] \\ & \leq \frac{c_f}{\min_{k \in \mathcal{D}_f} \{\sigma_k\}} \int_0^T \frac{|\kappa_H - N\kappa_H^N|}{2} \left(\mathbb{E} \left[\left\| \frac{1}{(H_{uN}^N)^2} \right\|_{\sigma} \right] + c \mathbb{E} \left[\left\| (H_{uN}^N)^2 \right\|_{\sigma} \right] \right) \\ & \quad + |\alpha - N\alpha^N| \|\mathbb{1}\|_{\sigma} + \frac{|\beta_H - N\beta_H^N|}{2} \mathbb{E} \left[\left\| \frac{1}{H_{uN}^N} \right\|_{\sigma} \right] du. \end{aligned} \quad (4.130)$$

Hence, from Lemmas 4.1.6 and 4.1.7 it follows for all $f \in \text{Dom}(\mathcal{A})$ and all $T \in [0, \infty)$ that

$$\begin{aligned}
 0 &\leq \lim_{N \rightarrow \infty} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \varepsilon_f^N(tN) \right| \right] \\
 &\leq \lim_{N \rightarrow \infty} \frac{T c_f}{\min_{k \in \mathcal{D}_f} \{\sigma_k\}} \left(\frac{|\kappa_H - N \kappa_H^N|}{2} \sup_{M \in \mathbb{N}} \sup_{t \in [0, \infty)} \left(\mathbb{E} \left[\left\| \frac{1}{(H_t^M)^2} \right\|_\sigma \right] \right. \right. \\
 &\quad \left. \left. + c \mathbb{E} \left[\left\| (H_t^M)^2 \right\|_\sigma \right] \right) + |\alpha - N \alpha^N| \|\underline{1}\|_\sigma + \frac{|\beta_H - N \beta_H^N|}{2} \sup_{M \in \mathbb{N}} \sup_{t \in [0, \infty)} \mathbb{E} \left[\left\| \frac{1}{H_t^M} \right\|_\sigma \right] \right) \\
 &= 0.
 \end{aligned} \tag{4.131}$$

Define the set

$$\begin{aligned}
 \mathcal{R} &:= \left\{ \bigtimes_{i \in \mathcal{D}} B_i : (B_i)_{i \in \mathcal{D}} \subseteq \mathcal{B}([0, \infty)^\mathcal{D}), \right. \\
 &\quad \left. B_i = [0, \infty) \text{ for all but finitely many } i \in \mathcal{D} \right\}.
 \end{aligned} \tag{4.132}$$

For all $N \in \mathbb{N}$, all $t \in [0, \infty)$, and all $B \in \mathcal{R}$ define the measure-valued random variables

$$\Lambda^N([0, t] \times B) := \int_0^t \mathbb{1}_B(H_{uN}^N) du = \int_0^t \prod_{i \in \mathcal{D}} \mathbb{1}_{B_i}(H_{uN}^N(i)) du, \tag{4.133}$$

Due to Carathéodory's theorem (e.g., Klenke, 2008, Theorem 1.41) there is a unique extension of this pre-measure to a measure on $[0, t] \times E_2$, which will be denoted by the same name. Define the space $\ell(E_2) := \{\mu : \mu \text{ is a measure on } [0, \infty) \times E_2 \text{ such that for all } t \in [0, \infty) \text{ it holds that } \mu([0, t] \times E_2) = t\}$ and define the space $D([0, \infty)) := \{f : [0, \infty) \rightarrow E_1 | f \text{ is càdlàg}\}$. Having checked all assumptions, Theorem 2.1 from Kurtz (1992) can now be applied, to show that the sequence $\{((F_{tN}^N)_{t \in [0, \infty)}, \Lambda^N) : N \in \mathbb{N}\}$ is relatively compact in $D([0, \infty)) \times \ell(E_2)$. Let (F, Λ) be a $D([0, \infty)) \times \ell(E_2)$ -valued random variable and let $(N_k)_{k \in \mathbb{N}} \subseteq \mathbb{N}$ be an increasing sequence such that $\lim_{k \rightarrow \infty} ((F_{tN_k}^{N_k})_{t \in [0, \infty)}, \Lambda^{N_k}) = (F, \Lambda)$. Due to Skorohod's representation theorem (see Ethier and Kurtz, 1986, Theorem 3.1.8) it can be assumed without loss of generality and for ease of notation that (F, Λ) acts on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Using Hölder's inequality and Theorem 3.1.1 it follows for all

$t \in [0, \infty)$ that

$$\begin{aligned}
 0 &\leq \lim_{N \rightarrow \infty} \int_0^t \mathbb{E} \left[\left\| H_{uN}^N - (h_\infty(F_{uN}^N(i)))_{i \in \mathcal{D}} \right\|_\sigma \right] du \\
 &= \lim_{N \rightarrow \infty} \int_0^t \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i |H_{uN}^N(i) - h_\infty(F_{uN}^N(i))| \right] du \\
 &\leq \lim_{N \rightarrow \infty} \sqrt{\int_0^t \mathbb{E} \left[\sum_{i \in \mathcal{D}} \sigma_i (H_{uN}^N(i) - h_\infty(F_{uN}^N(i)))^2 \right] du} \sqrt{t \sum_{k \in \mathcal{D}} \sigma_k} = 0.
 \end{aligned} \tag{4.134}$$

For any bounded Lipschitz continuous function $f: l_\sigma^1 \rightarrow \mathbb{R}$ with Lipschitz constant \bar{c}_f and all $t \in [0, \infty)$, applying (4.134), it follows that

$$\begin{aligned}
 0 &\leq \mathbb{E} \left[\left| \int_0^t \int_{E_2} f(y) \Lambda(du \times dy) - \int_0^t f(h_\infty(F_u)) du \right| \right] \\
 &= \lim_{k \rightarrow \infty} \mathbb{E} \left[\left| \int_0^t f(H_{uN_k}^{N_k}) du - \int_0^t f(h_\infty(F_{uN_k}^{N_k})) du \right| \right] \\
 &\leq \bar{c}_f \lim_{k \rightarrow \infty} \mathbb{E} \left[\int_0^t \|H_{uN_k}^{N_k} - h_\infty(F_{uN_k}^{N_k})\| du \right] = 0.
 \end{aligned} \tag{4.135}$$

Define the operator $\mathcal{A}_2: \text{Dom}(\mathcal{A}) \rightarrow C(E_1, \mathbb{R})$ for all $f \in \text{Dom}(\mathcal{A})$ and all $x \in E_1$ by

$$\begin{aligned}
 (\mathcal{A}_2 f)(x) &:= \sum_{i \in \mathcal{D}} \left(\kappa_H \sum_{j \in \mathcal{D}} \left(m(i, j) \frac{a-x_i}{a-x_j} (x_j - x_i) \right) - \alpha x_i (1 - x_i) \right) \frac{df}{dx_i}(x) \\
 &\quad + \frac{1}{2} \sum_{i \in \mathcal{D}} \beta_H b(a - x_i) x_i (1 - x_i) \frac{d^2 f}{dx_i^2}(x).
 \end{aligned} \tag{4.136}$$

For all $t \in [0, \infty)$, all $f \in \text{Dom}(\mathcal{A})$, and all $x \in E_1$ it holds \mathbb{P} -a.s. that

$$\begin{aligned}
 &\int_0^t \int_{E_2} (\mathcal{A}_1 f)(F_s, y) \Lambda(ds \times dy) \\
 &= \int_0^t \int_{E_2} \left[\sum_{i \in \mathcal{D}} \left(\kappa_H \sum_{j \in \mathcal{D}} \left(m(i, j) \frac{y_j}{y_i} (F_s(j) - F_s(i)) \right) - \alpha F_s(i) (1 - F_s(i)) \right) \frac{df}{dx_i}(F_s) \right. \\
 &\quad \left. + \frac{1}{2} \sum_{i \in \mathcal{D}} \beta_H \frac{F_s(i)(1-F_s(i))}{y_i} \frac{d^2 f}{dx_i^2}(F_s) \right] \prod_{i \in \mathcal{D}} \mathbb{1}_{y_i}(h_\infty(F_s(i))) dy ds \\
 &= \int_0^t (\mathcal{A}_2 f)(F_s) ds.
 \end{aligned} \tag{4.137}$$

Applying Theorem 2.1 of Kurtz (1992) together with (4.137), it follows for each $f \in$

$\text{Dom}(\mathcal{A})$ that

$$\left(f(F_t) - \int_0^t (\mathcal{A}_2 f)(F_s) ds\right)_{t \in [0, \infty)} \quad (4.138)$$

is a martingale. Hence, F is a solution to (3.2). Note that for all $z_1, z_2 \in [0, 1]$ it holds that $\frac{a-z_1}{a-z_2}(z_2 - z_1) = (a - z_1)(\frac{a-z_1}{a-z_2} - 1)$. Using this and (4.10) it then follows for any subset $\mathcal{S} \subseteq \mathcal{D}$ and any $x, y \in E_2$ that

$$\begin{aligned} & \sum_{i \in \mathcal{S}} \sigma_i \mathbb{1}_{x_i \geq y_i} \left(\kappa_H \sum_{j \in \mathcal{D}} m(i, j) \left(\frac{(a-x_i)^2}{a-x_j} - (a-x_i) - \frac{(a-y_i)^2}{a-y_j} + (a-y_i) \right) \right. \\ & \quad \left. - \alpha(x_i(1-x_i) - y_i(1-y_i)) \right) \\ &= \sum_{i \in \mathcal{S}} \sigma_i \mathbb{1}_{x_i \geq y_i} \left(\kappa_H \sum_{j \in \mathcal{D}} m(i, j) \left((x_i - y_i) + ((a-x_i)^2 - (a-y_i)^2) \frac{1}{a-x_j} \right. \right. \\ & \quad \left. \left. - (a-y_i)^2 \left(\frac{1}{a-y_j} - \frac{1}{a-x_j} \right) \right) + \alpha(-(x_i - y_i) + x_i^2 - y_i^2) \right) \\ &\leq \sum_{i \in \mathcal{S}} \sigma_i \left(\kappa_H \sum_{j \in \mathcal{D}} m(i, j) \mathbb{1}_{x_j \geq y_j} (a-y_i)^2 \left(\frac{1}{a-x_j} - \frac{1}{a-y_j} \right) \right) + \sum_{i \in \mathcal{S}} \sigma_i (\kappa_H \\ & \quad + 2\alpha) \mathbb{1}_{x_i \geq y_i} (x_i - y_i) \\ &\leq \sum_{i \in \mathcal{S}} \sigma_i \left(\kappa_H \sum_{j \in \mathcal{D}} m(i, j) \mathbb{1}_{x_j \geq y_j} \frac{a^2}{(a-1)^2} (x_j - y_j) \right) + \sum_{i \in \mathcal{S}} \sigma_i (\kappa_H \\ & \quad + 2\alpha) \mathbb{1}_{x_i \geq y_i} (x_i - y_i) \\ &\leq \sum_{i \in \mathcal{S}} \sigma_i c \kappa_H \mathbb{1}_{x_i \geq y_i} \frac{a^2(x_i - y_i)}{(a-1)^2} + \sum_{i \in \mathcal{S}} \sigma_i (\kappa_H + 2\alpha) \mathbb{1}_{x_i \geq y_i} (x_i - y_i) \\ &= \sum_{i \in \mathcal{S}} \sigma_i \left(\frac{c \kappa_H a^2}{(a-1)^2} + \kappa_H + 2\alpha \right) (x_i - y_i)^+. \end{aligned} \quad (4.139)$$

This implies that equation (26) of Hutzenthaler and Wakolbinger (2007) is fulfilled. Together with the assumptions on m in Assumption 4.1.1 it now follows, analogous to Proposition 2.1 of Hutzenthaler and Wakolbinger (2007), that the system (3.2) has a unique strong solution with a.s. continuous paths. Consequently, any limit point of $\{ (F_{tN}^N)_{t \in [0, \infty)} : N \in \mathbb{N} \}$ solves (3.2). Combining this with the fact that $\{ (F_{tN}^N)_{t \in [0, \infty)} : N \in \mathbb{N} \}$ is relatively compact it follows that $(F_{tN}^N)_{t \in [0, \infty)} \Rightarrow (X_t)_{t \in [0, \infty)}$, as $N \rightarrow \infty$. This completes the proof of Theorem 3.1.2. \square

4.1.5 McKean-Vlasov limit

In this section the convergence of a sequence of exchangeable systems of stochastic differential equations is investigated in the limit of infinitely many components. This result is then applied to the case of altruistic defense traits and the many-demes limit is derived.

4.1.5.1 Setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $I \subset [0, \infty)$ be an interval of length $|I| \in (0, \infty]$ which is either of the form $[0, |I|]$ if $|I| < \infty$ or of the form $[0, \infty)$ if $|I| = \infty$, let $A \subseteq \mathbb{R}$ be a convex set, and let $\psi: I \rightarrow A$, $\xi: A \times I \rightarrow \mathbb{R}$, and $\sigma^2: I \rightarrow [0, \infty)$ be functions. The function $\sigma^2: I \rightarrow [0, \infty)$ is locally Lipschitz continuous in I and satisfies $\sigma^2(0) = 0$ and if $|I| < \infty$, then $\sigma^2(|I|) = 0$. Furthermore, the function σ^2 is strictly positive on $(0, |I|)$. There exists a constant $L \in (0, \infty)$ such that σ^2 satisfies the growth condition that for all $y \in I$ it holds that $\sigma^2(y) \leq L(y + y^2)$ and such that ξ satisfies for all $(u, x), (v, y) \in A \times I$ that

$$\mathbb{1}_{x \geq y} (\xi(u, x) - \xi(v, y)) \leq L(u - v)^+ + L(x - y)^+. \quad (4.140)$$

The function $\psi: I \rightarrow [0, \infty)$ satisfies for all $x, y \in I$ that $|\psi(x) - \psi(y)| \leq L|x - y|$. Let $W(i): [0, \infty) \times \Omega \rightarrow \mathbb{R}$, $i \in \mathbb{N}$, be independent Brownian motions with continuous sample paths. For all $D \in \mathbb{N}$ let $X^D: [0, \infty) \times \{1, \dots, D\} \times \Omega \rightarrow I$ be an adapted stochastic process with continuous sample paths that for all $t \in [0, \infty)$ and all $i \in \{1, \dots, D\}$ \mathbb{P} -a.s. satisfies

$$\begin{aligned} X_t^D(i) &= X_0^D(i) + \int_0^t \xi\left(\frac{1}{D} \sum_{j \in \{1, \dots, D\}} \psi(X_s^D(j)), X_s^D(i)\right) ds \\ &\quad + \int_0^t \sqrt{\sigma^2(X_s^D(i))} dW_s(i). \end{aligned} \quad (4.141)$$

Let $M: [0, \infty) \times \Omega \rightarrow I$ be an adapted stochastic process with continuous sample paths that for all $t \in [0, \infty)$ \mathbb{P} -a.s. satisfies

$$M_t = M_0 + \int_0^t \xi(\mathbb{E}[\psi(M_s)], M_s) ds + \int_0^t \sqrt{\sigma^2(M_s)} dW_s(1). \quad (4.142)$$

4.1.5.2 McKean-Vlasov limit

The following proposition, Proposition 4.1.13, partly generalizes Proposition 4.29 of Hutzenhaler (2012) where ξ depends linearly on its first argument.

Proposition 4.1.13. *Assume the setting of Section 4.1.5.1, let M_0 be an I -valued random variable, for every $D \in \mathbb{N}$ let $(X_0^D(j))_{j \in \{1, \dots, D\}}$ be exchangeable and integrable random variables with values in I . Then, there exists a unique solution M of (4.142) and for all*

$D \in \mathbb{N}$ and all $t \in [0, \infty)$ it holds that

$$\begin{aligned} & \sqrt{D} \mathbb{E} [|X_t^D(1) - M_t|] \\ & \leq e^{(L^2 + L + L_\mu)t} \left(\sqrt{D} \mathbb{E} [|X_0^D(1) - M_0|] + L \int_0^t \left(\text{Var}(\psi(M_s)) \right)^{\frac{1}{2}} ds \right). \end{aligned} \quad (4.143)$$

Proof. Existence of a weak solution is straightforward using a tightness argument. Next, pathwise uniqueness for the SDE (4.142) is shown. Let $M, \bar{M}: [0, \infty) \times \Omega \rightarrow I$ be two solutions of the SDE (4.142). Then our assumptions and a standard Yamada-Watanabe argument (e.g., Yamada and Watanabe, 1971, Theorem 1) shows for all $t \in [0, \infty)$ that \mathbb{P} -a.s.

$$|M_t - \bar{M}_t| = |M_0 - \bar{M}_0| + \int_0^t \text{sgn}(M_s - \bar{M}_s) d(M_s - \bar{M}_s). \quad (4.144)$$

Let $(\tau_l)_{l \in \mathbb{N}}$ be a localizing sequence for the local martingale $(\int_0^t \text{sgn}(M_s - \bar{M}_s)(\sigma^2(M_s) - \sigma^2(\bar{M}_s)) dW_s)_{t \in [0, \infty)}$. Then Fatou's Lemma and our assumptions imply for all $t \in [0, \infty)$ that

$$\begin{aligned} \mathbb{E}[|M_t - \bar{M}_t|] & \leq \lim_{l \rightarrow \infty} \mathbb{E}[|M_{t \wedge \tau_l} - \bar{M}_{t \wedge \tau_l}|] \leq \mathbb{E}[|M_0 - \bar{M}_0|] \\ & \quad + \mathbb{E} \left[\int_0^t \text{sgn}(M_s - \bar{M}_s) (\xi(\mathbb{E}[\psi(M_s)], M_s) - \xi(\mathbb{E}[\psi(\bar{M}_s)], \bar{M}_s)) ds \right] \\ & \leq \mathbb{E}[|M_0 - \bar{M}_0|] + L \int_0^t |\mathbb{E}[\psi(M_s)] - \mathbb{E}[\psi(\bar{M}_s)]| + \mathbb{E}[|M_s - \bar{M}_s|] ds \\ & \leq \mathbb{E}[|M_0 - \bar{M}_0|] + (L + 1)^2 \int_0^t \mathbb{E}[|M_s - \bar{M}_s|] ds. \end{aligned} \quad (4.145)$$

This together with Gronwall's lemma implies pathwise uniqueness for the SDE (4.142). Therefore, the theorem of Yamada and Watanabe (1971) implies that the SDE (4.142) is exact. The rest of the proof is analogous to the proof of Proposition 4.29 of Hutzenthaler (2012) and it is omitted here. \square

4.1.5.3 Application to altruistic defense in structured populations

In this section, the applicability of Proposition 4.1.13 to the case of altruistic defense in structured populations is verified.

Lemma 4.1.14. *Let $\alpha, \beta, \kappa \in (0, \infty)$ and $a \in (1, \infty)$, let $I = [0, 1]$ and define the function $\sigma^2: I \rightarrow [0, \infty)$ by $I \ni x \mapsto \sigma^2(x) := \beta(a - x)x(1 - x)$, the function $\psi: I \rightarrow [0, \infty)$ by $I \ni x \mapsto \psi(x) := \frac{1}{a-x}$, and the function $\xi: [0, \infty) \times I \rightarrow \mathbb{R}$ by $[0, \infty) \times I \ni (u, x) \mapsto$*

$\xi(u, x) := \kappa(a - x)((a - x)u - 1) - \alpha x(1 - x)$. Then the interval I and the functions σ^2 , ψ , and ξ satisfy the setting of Section 4.1.5.1 with $L = \max\{\beta a, \kappa a^2, \kappa + \alpha, \frac{1}{(a-1)^2}\}$.

Proof. For all $(u, x), (v, y) \in [0, \infty) \times [0, 1]$ it holds that

$$\begin{aligned}
 \mathbb{1}_{x \geq y}(\xi(u, x) - \xi(v, y)) &= \mathbb{1}_{x \geq y}(\kappa(a - x)((a - x)u - 1) \\
 &\quad - \kappa(a - y)((a - y)v - 1) - \alpha x(1 - x) + \alpha y(1 - y)) \\
 &= \mathbb{1}_{x \geq y}(\kappa[(a - x)^2 u - (a - x) - (a - y)^2 v + (a - y)] \\
 &\quad - \alpha(1 - (x + y))(x - y)) \\
 &= \mathbb{1}_{x \geq y}(\kappa[(x - y) + ((a - x)^2 - (a - y)^2)u \\
 &\quad - (a - y)^2(v - u)] - \alpha(1 - (x + y))(x - y)) \\
 &\leq (\kappa + \alpha)(x - y)^+ + \kappa a^2(u - v)^+ \leq L(x - y)^+ + L(u - v)^+.
 \end{aligned} \tag{4.146}$$

Moreover, for all $x, y \in I$ it holds that $\sigma^2(x) = \beta(a - x)x(1 - x) \leq \beta a x \leq L(x + x^2)$ and that

$$|\psi(x) - \psi(y)| = \left| \frac{1}{a-x} - \frac{1}{a-y} \right| = \left| \int_y^x \frac{1}{(a-z)^2} dz \right| \leq \frac{1}{(a-1)^2} |x - y| \leq L|x - y|. \tag{4.147}$$

This completes the proof of Lemma 4.1.14. \square

4.1.6 Long-term behavior of the average altruist frequency

In this section, the many-demes limit of the process describing the altruist frequency is investigated in the limit $t \rightarrow \infty$.

4.1.6.1 Setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\kappa, \alpha, \beta \in (0, \infty)$, $a \in (1, \infty)$, $c \in (0, 1)$, let $W: [0, \infty) \times \Omega \rightarrow \mathbb{R}$ be a Brownian motion with continuous sample paths, let $Z: [0, \infty) \times \Omega \rightarrow [0, 1]$ be an adapted process with continuous sample paths that for all $t \in [0, \infty)$ satisfies \mathbb{P} -a.s.

$$\begin{aligned}
 Z_t &= Z_0 + \int_0^t (\kappa(a - Z_s)((a - Z_s)\mathbb{E}\left[\frac{1}{a-Z_s}\right] - 1) - \alpha Z_s(1 - Z_s)) ds \\
 &\quad + \int_0^t \sqrt{\beta(a - Z_s)Z_s(1 - Z_s)} dW_s.
 \end{aligned} \tag{4.148}$$

Moreover, for all $\theta \in (\frac{1}{a}, \frac{1}{a-1})$ let $Z^\theta: [0, \infty) \times \Omega \rightarrow [0, 1]$ be an adapted process with continuous sample paths that for all $t \in [0, \infty)$ satisfies \mathbb{P} -a.s.

$$\begin{aligned} Z_t^\theta &= Z_0^\theta + \int_0^t (\kappa(a - Z_s^\theta) ((a - Z_s^\theta)\theta - 1) - \alpha Z_s^\theta (1 - Z_s^\theta)) ds \\ &\quad + \int_0^t \sqrt{\beta(a - Z_s^\theta) Z_s^\theta (1 - Z_s^\theta)} dW_s. \end{aligned} \quad (4.149)$$

For all $\theta \in (\frac{1}{a}, \frac{1}{a-1})$ and all $z \in [0, 1]$ define

$$\begin{aligned} m_\theta(z) &:= \beta c^{\frac{2\kappa}{\beta}(a\theta-1)} (1-c)^{\frac{2\kappa}{\beta}(1-\theta(a-1))} (a-c)^{\frac{2\alpha}{\beta} \frac{1}{\beta(a-z)z(1-z)}} \\ &\quad \cdot \exp\left(\int_c^z 2 \frac{\kappa(a-y)((a-y)\theta-1)-\alpha y(1-y)}{\beta(a-y)y(1-y)} dy\right) \\ &= z^{\frac{2\kappa}{\beta}(a\theta-1)-1} (1-z)^{\frac{2\kappa}{\beta}(1-\theta(a-1))-1} (a-z)^{\frac{2\alpha}{\beta}-1}. \end{aligned} \quad (4.150)$$

Note that this defines the speed density (see Karlin and Taylor, 1981, p. 95) for (4.149). Furthermore, note that for all $\theta \in (\frac{1}{a}, \frac{1}{a-1})$ it holds that

$$\int_0^1 m_\theta(z) dz < \infty. \quad (4.151)$$

For all $\theta \in (\frac{1}{a}, \frac{1}{a-1})$ define $c_\theta := \int_0^1 m_\theta(z) dz$, for all $x \in \{0, 1\}$ denote by δ_x the Dirac measure on $[0, 1]$, and for all $\theta \in [\frac{1}{a}, \frac{1}{a-1}]$ define the mapping $\Psi_\theta: \mathcal{B}([0, 1]) \rightarrow [0, 1]$ by

$$\mathcal{B}([0, 1]) \ni A \mapsto \Psi_\theta(A) := \begin{cases} \delta_0(A), & \text{if } \theta = \frac{1}{a}, \\ \delta_1(A), & \text{if } \theta = \frac{1}{a-1}, \\ \int_A \frac{1}{c_\theta} m_\theta(z) dz, & \text{if } \theta \in (\frac{1}{a}, \frac{1}{a-1}). \end{cases} \quad (4.152)$$

4.1.6.2 Results for the equilibrium distribution

Assume the setting of Section 4.1.6.1. Existence and uniqueness of the solution of (4.148) follow from Proposition 4.1.13. When $\theta \in (\frac{1}{a}, \frac{1}{a-1})$ it holds that Ψ_θ defines a probability distribution by (4.151), and Theorem V.54.5 of Rogers and Williams (2000) can be applied to conclude that it is the unique equilibrium distribution for (4.149). The proof of the following lemma, Lemma 4.1.15, is clear and therefore omitted.

Lemma 4.1.15. *Assume the setting of Section 4.1.6.1. A probability measure $\Phi: \mathcal{B}([0, 1]) \rightarrow [0, 1]$ is an equilibrium distribution of the dynamics (4.148) if and only if there exists a $\theta \in [\frac{1}{a}, \frac{1}{a-1}]$ such that $\Phi = \Psi_\theta$.*

Lemma 4.1.16. Assume the setting of Section 4.1.6.1 and let $\theta \in (\frac{1}{a}, \frac{1}{a-1})$. Then it holds that

$$\int_0^1 \frac{1}{a-z} \Psi_\theta(dz) \begin{cases} < \theta, \text{ if } \alpha > \beta, \\ = \theta, \text{ if } \alpha = \beta, \\ > \theta, \text{ if } \alpha < \beta. \end{cases} \quad (4.153)$$

Proof. Define $u := \frac{2\kappa}{\beta}(a\theta - 1)$ and $v := \frac{2\kappa}{\beta}(1 - \theta(a - 1))$ and note that $u, v \in (0, \infty)$. Let $\Gamma: (0, \infty) \rightarrow (0, \infty)$ be the Gamma function, i.e., for all $x \in (0, \infty)$ let $\Gamma(x) := \int_0^\infty z^{x-1} e^{-z} dz$. It is well-known that for all $x \in (0, \infty)$ the Gamma function satisfies $\Gamma(x+1) = x\Gamma(x)$ and that for all $x, y \in (0, \infty)$ it holds that $\int_0^1 z^{x-1}(1-z)^{y-1} dz = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$. Thus, it follows that

$$\begin{aligned} & \int_0^1 z^{u-1}(1-z)^{v-1}(a-z) \left(\frac{1}{a-z} - \theta \right) dz \\ &= \int_0^1 z^{u-1}(1-z)^{v-1} dz - a\theta \int_0^1 z^{u-1}(1-z)^{v-1} dz + \theta \int_0^1 z^u(1-z)^{v-1} dz \\ &= \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)} - a\theta \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)} + \theta \frac{\Gamma(u+1)\Gamma(v)}{\Gamma(u+v+1)} = \left((1-a\theta) \frac{(u+v)\Gamma(u)\Gamma(v)}{\Gamma(u+v+1)} + \theta \frac{u\Gamma(u)\Gamma(v)}{\Gamma(u+v+1)} \right) \\ &= (u(1-a\theta) + v(1-a\theta)) \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v+1)} \\ &= \frac{2\kappa}{\beta} ((a\theta - 1)(1 - \theta(a - 1)) + (1 - \theta(a - 1))(1 - a\theta)) \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v+1)} \\ &= \left(\frac{2\kappa}{\beta} (1 - \theta(a - 1))(a\theta - 1 + 1 - a\theta) \right) \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v+1)} = 0. \end{aligned} \quad (4.154)$$

First, consider the case $\alpha = \beta$. Using (4.154) it follows that

$$\begin{aligned} \int_0^1 \frac{1}{a-z} \Psi_\theta(dz) - \theta &= \int_0^1 c_\theta z^{\frac{2\kappa}{\beta}(a\theta-1)-1} (1-z)^{\frac{2\kappa}{\beta}(1-\theta(a-1))-1} (a-z)^{\frac{2\alpha}{\beta}-1} \left(\frac{1}{a-z} \right. \\ &\quad \left. - \theta \right) dz = c_\theta \int_0^1 z^{u-1}(1-z)^{v-1}(a-z) \left(\frac{1}{a-z} - \theta \right) dz = 0. \end{aligned} \quad (4.155)$$

Now, consider the case $\alpha > \beta$. Let $\hat{\delta} := \alpha - \beta$, $\delta := \frac{2\hat{\delta}}{\beta}$, and $z^* := \sup\{z \in (0, 1) : \frac{1}{a-z} - \theta < 0\}$. Note that $\hat{\delta}, \delta > 0$ and $z^* = a - \frac{1}{\theta} \in (0, 1)$. Also note that for all $z \in (0, z^*)$ it holds that $\frac{1}{a-z} - \theta < 0$ and $(a-z)^\delta > (a-z^*)^\delta$. Furthermore, for all $z \in (z^*, 1)$ it holds

that $\frac{1}{a-z} - \theta > 0$ and $(a-z)^\delta < (a-z^*)^\delta$. Together with (4.154) it thereby follows that

$$\begin{aligned}
 & \int_0^1 \frac{1}{a-z} \Psi_\theta(dz) - \theta = \int_0^1 \left(\frac{1}{a-z} - \theta \right) \Psi_\theta(dz) \\
 & = \int_0^1 c_\theta z^{u-1} (1-z)^{v-1} (a-z)^{\frac{2\alpha}{\beta}-1} \left(\frac{1}{a-z} - \theta \right) dz \\
 & = \int_0^{z^*} c_\theta z^{u-1} (1-z)^{v-1} (a-z)^{1+\delta} \left(\frac{1}{a-z} - \theta \right) dz \\
 & \quad + \int_{z^*}^1 c_\theta z^{u-1} (1-z)^{v-1} (a-z)^{1+\delta} \left(\frac{1}{a-z} - \theta \right) dz \\
 & < c_\theta (a-z^*)^\delta \left(\int_0^{z^*} z^{u-1} (1-z)^{v-1} (a-z) \left(\frac{1}{a-z} - \theta \right) dz \right. \\
 & \quad \left. + \int_{z^*}^1 z^{u-1} (1-z)^{v-1} (a-z) \left(\frac{1}{a-z} - \theta \right) dz \right) \\
 & = c_\theta (a-z^*)^\delta \int_0^1 z^{u-1} (1-z)^{v-1} (a-z) \left(\frac{1}{a-z} - \theta \right) dz = 0.
 \end{aligned} \tag{4.156}$$

The case $\alpha < \beta$ can be proved analogously and is omitted here. This completes the proof. \square

4.1.6.3 Proof of Theorem 3.1.3

Proof of Theorem 3.1.3. Applying Itô's lemma yields for all $t \in [0, \infty)$ that

$$\begin{aligned}
 \frac{1}{a-Z_t} - \frac{1}{a-Z_0} &= \int_0^t \frac{1}{(a-Z_s)^2} \left(\kappa(a-Z_s) \left((a-Z_s) \mathbb{E} \left[\frac{1}{a-Z_s} \right] - 1 \right) \right. \\
 &\quad \left. - \alpha Z_s (1-Z_s) \right) + \frac{1}{2} \frac{2(a-Z_s)}{(a-Z_s)^4} \beta (a-Z_s) Z_s (1-Z_s) ds \\
 &+ \int_0^t \frac{1}{(a-Z_s)^2} \sqrt{\beta(a-Z_s) Z_s (1-Z_s)} dW_s = \int_0^t \kappa \left(\mathbb{E} \left[\frac{1}{a-Z_s} \right] - \frac{1}{a-Z_s} \right) \\
 &\quad - \frac{\alpha Z_s (1-Z_s)}{(a-Z_s)^2} + \frac{\beta Z_s (1-Z_s)}{(a-Z_s)^2} ds + \int_0^t \frac{1}{(a-Z_s)^2} \sqrt{\beta(a-Z_s) Z_s (1-Z_s)} dW_s.
 \end{aligned} \tag{4.157}$$

After taking expectations, Fubini's theorem yields for all $t \in [0, \infty)$ that

$$\begin{aligned}
 \mathbb{E} \left[\frac{1}{a-Z_t} \right] - \mathbb{E} \left[\frac{1}{a-Z_0} \right] &= \int_0^t \kappa \left(\mathbb{E} \left[\frac{1}{a-Z_s} \right] - \mathbb{E} \left[\frac{1}{a-Z_s} \right] \right) - \alpha \mathbb{E} \left[\frac{Z_s (1-Z_s)}{(a-Z_s)^2} \right] \\
 &\quad + \beta \mathbb{E} \left[\frac{Z_s (1-Z_s)}{(a-Z_s)^2} \right] ds = (\beta - \alpha) \int_0^t \mathbb{E} \left[\frac{Z_s (1-Z_s)}{(a-Z_s)^2} \right] ds.
 \end{aligned} \tag{4.158}$$

Since for all $s \in [0, \infty)$ it holds that $\mathbb{E}\left[\frac{Z_s(1-Z_s)}{(a-Z_s)^2}\right] \geq 0$ it follows that the function $[0, \infty) \ni t \mapsto \mathbb{E}\left[\frac{1}{a-Z_t}\right] \in \left[\frac{1}{a}, \frac{1}{a-1}\right]$ converges monotonically non-increasing as $t \rightarrow \infty$ if $\alpha > \beta$, monotonically non-decreasing if $\alpha < \beta$, or is constant if $\alpha = \beta$.

First, assume $\alpha > \beta$. From (4.148) it follows that δ_1 is an invariant measure for Z . So if $\mathbb{P}[Z_0 = 1] = 1$, then for all $t \in [0, \infty)$ it holds that $\mathbb{P}[Z_t = 1] = 1$. Now let $\mathbb{P}[Z_0 = 1] < 1$, implying $\mathbb{E}\left[\frac{1}{a-Z_0}\right] \in \left[\frac{1}{a}, \frac{1}{a-1}\right)$. Define $\theta := \lim_{t \rightarrow \infty} \mathbb{E}\left[\frac{1}{a-Z_t}\right]$ and fix it for the rest of the paragraph. Note that due to the monotonicity stated above it holds that $\theta \in \left[\frac{1}{a}, \frac{1}{a-1}\right)$. Aiming at a contradiction, assume that $\theta \in \left(\frac{1}{a}, \frac{1}{a-1}\right)$. Choose any $\varepsilon \in \left(0, \frac{1}{a-1} - \theta\right)$ and fix it for the rest of the proof. By definition of θ there exists an $s_\varepsilon \in (0, \infty)$, such that for all $t \in [s_\varepsilon, \infty)$ it holds that $\mathbb{E}\left[\frac{1}{a-Z_t}\right] < \theta + \varepsilon$. Let $\tilde{W} : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ be a Brownian motion with continuous sample paths, let $\tilde{Z} : [0, \infty) \times \Omega \rightarrow [0, 1]$ and $\tilde{Z}^{\theta+\varepsilon} : [0, \infty) \times \Omega \rightarrow [0, 1]$ be adapted processes with continuous sample paths that satisfy for all $t \in [0, \infty)$ \mathbb{P} -a.s.

$$\begin{aligned} \tilde{Z}_t &= \tilde{Z}_0 + \int_0^t \left(\kappa(a - \tilde{Z}_s) \left((a - \tilde{Z}_s) \mathbb{E}\left[\frac{1}{a-\tilde{Z}_s}\right] - 1 \right) - \alpha \tilde{Z}_s(1 - \tilde{Z}_s) \right) ds \\ &\quad + \int_0^t \sqrt{\beta(a - \tilde{Z}_s)\tilde{Z}_s(1 - \tilde{Z}_s)} d\tilde{W}_s, \\ \tilde{Z}_t^{\theta+\varepsilon} &= \tilde{Z}_0^{\theta+\varepsilon} + \int_0^t \left(\kappa(a - \tilde{Z}_s^{\theta+\varepsilon}) \left((a - \tilde{Z}_s^{\theta+\varepsilon})(\theta + \varepsilon) - 1 \right) \right. \\ &\quad \left. - \alpha \tilde{Z}_s^{\theta+\varepsilon}(1 - \tilde{Z}_s^{\theta+\varepsilon}) \right) ds + \int_0^t \sqrt{\beta(a - \tilde{Z}_s^{\theta+\varepsilon})\tilde{Z}_s^{\theta+\varepsilon}(1 - \tilde{Z}_s^{\theta+\varepsilon})} d\tilde{W}_s, \end{aligned} \quad (4.159)$$

such that $\tilde{Z}_0^{\theta+\varepsilon} = \tilde{Z}_0$ and such that \tilde{Z}_0 and Z_{s_ε} are equal in distribution. Then for each $t \in [s_\varepsilon, \infty)$ it holds that Z_t and $\tilde{Z}_{t-s_\varepsilon}$ are equal in distribution and the drift term of $\tilde{Z}_{t-s_\varepsilon}$ is lower than that of $\tilde{Z}_{t-s_\varepsilon}^{\theta+\varepsilon}$. Together with the fact that the mapping $[0, 1] \ni z \mapsto \frac{1}{a-z}$ is strictly monotonically increasing this implies for all $t \in [s_\varepsilon, \infty)$ that

$$\mathbb{E}\left[\frac{1}{a-Z_t}\right] = \mathbb{E}\left[\frac{1}{a-\tilde{Z}_{t-s_\varepsilon}}\right] \leq \mathbb{E}\left[\frac{1}{a-\tilde{Z}_{t-s_\varepsilon}^{\theta+\varepsilon}}\right]. \quad (4.160)$$

Recall from Section 4.1.6.2 that for any $\eta \in \left(\frac{1}{a}, \frac{1}{a-1}\right)$ it holds that Ψ_η is the unique equilibrium distribution of \tilde{Z}^η . Combining this with (4.160), it follows that (cf., Rogers and Williams, 2000, Theorem V.54.5)

$$\theta = \lim_{t \rightarrow \infty} \mathbb{E}\left[\frac{1}{a-Z_t}\right] \leq \lim_{t \rightarrow \infty} \mathbb{E}\left[\frac{1}{a-\tilde{Z}_{t-s_\varepsilon}^{\theta+\varepsilon}}\right] = \int_0^1 \frac{1}{a-z} \Psi_{\theta+\varepsilon}(dz). \quad (4.161)$$

The dominated convergence theorem yields that the mapping $\left(\frac{1}{a}, \frac{1}{a-1}\right) \ni \eta \mapsto \Psi_\eta$ is continuous with respect to the weak topology. Applying this, (4.161) together with the fact that

$\varepsilon \in (0, \frac{1}{a-1} - \theta)$ was arbitrarily chosen, and Lemma 4.1.16, yields the contradiction

$$\theta \leq \lim_{\delta \rightarrow 0} \int_0^1 \frac{1}{a-z} \Psi_{\theta+\delta}(dz) = \int_0^1 \frac{1}{a-z} \Psi_{\theta}(dz) < \theta. \quad (4.162)$$

Hence, it holds that $\theta = \frac{1}{a}$, implying

$$0 \leq \lim_{t \rightarrow \infty} \mathbb{E}[Z_t] \leq \lim_{t \rightarrow \infty} a^2 \mathbb{E}\left[\frac{Z_t}{a-Z_t}\right] = \lim_{t \rightarrow \infty} a^2 \mathbb{E}\left[\frac{1}{a-Z_t}\right] - a^2 \frac{1}{a} = 0. \quad (4.163)$$

The case $\alpha < \beta$ can be proved analogously and is omitted here.

Finally, assume $\alpha = \beta$, define $\theta := \mathbb{E}[\frac{1}{a-Z_0}]$, and fix it for the rest of the proof. From (4.158) it follows that $\mathbb{E}[\frac{1}{a-Z_t}]$ is constant in $t \in [0, \infty)$. Thus, assuming that Z_0 and Z_0^θ are equal in distribution it follows from (4.148) and (4.149) that for all $t \in [0, \infty)$ it holds that Z_t and Z_t^θ are equal in distribution. Recall from Section 4.1.6.2 that Ψ_θ is the unique equilibrium distribution of Z^θ . Consequently, Ψ_θ is the unique equilibrium distribution of Z . This completes the proof of Theorem 3.1.3. \square

4.1.7 Invasion of an altruistic defense allele

In this section, the invasion of an allele for altruistic defense in a structured population of non-altruistic individuals is investigated.

4.1.7.1 Setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\kappa, \alpha, \beta \in (0, \infty)$, $a \in (1, \infty)$, and let $W(i): [0, \infty) \times \Omega \rightarrow \mathbb{R}$, $i \in \mathbb{N}$, be independent Brownian motions with continuous sample paths. For all $D \in \mathbb{N}$ let $X^D: [0, \infty) \times \{1, \dots, D\} \times \Omega \rightarrow [0, 1]$ be an adapted process with continuous sample paths that for all $t \in [0, \infty)$ and all $i \in \{1, \dots, D\}$ \mathbb{P} -a.s. satisfies

$$\begin{aligned} X_t^D(i) = & X_0^D(i) + \int_0^t \kappa(a - X_s^D(i)) \left((a - X_s^D(i))^{\frac{1}{D}} \sum_{j=1}^D \frac{1}{a - X_s^D(j)} - 1 \right) \\ & - \alpha X_s^D(i)(1 - X_s^D(i)) ds + \int_0^t \sqrt{\beta(a - X_s^D(i))X_s^D(i)(1 - X_s^D(i))} dW_s(i). \end{aligned} \quad (4.164)$$

Let $\tilde{a}: [0, \infty) \rightarrow [0, \infty)$ be a function defined by

$$[0, \infty) \ni x \mapsto \tilde{a}(x) := \kappa a \frac{\min\{x, 1\}}{a - \min\{x, 1\}} + (x - 1)^+. \quad (4.165)$$

Then, assuming there is positive mass only in deme 1, the dynamics in deme 1 follows asymptotically the following process Y . Let $Y: [0, \infty) \times \Omega \rightarrow [0, 1]$ be an adapted process with continuous sample paths such that for all $t \in [0, \infty)$ it \mathbb{P} -a.s. holds that

$$Y_t = Y_0 - \int_0^t \frac{\kappa}{a} Y_s(a - Y_s) + \alpha Y_s(1 - Y_s) ds + \int_0^t \sqrt{\beta(a - Y_s)Y_s(1 - Y_s)} dW_s(1). \quad (4.166)$$

In addition, let Q_Y be the excursion measure which satisfies $Q_Y = \lim_{0 < \varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{P}[Y \in \cdot | Y_0 = \varepsilon]$ in a suitable sense (see Pitman and Yor, 1982; Hutzenthaler, 2009, for details). Asymptotically in the many-demes limit, every deme with population path $\chi \in C([0, \infty), [0, 1])$ populates demes through migration and these new populations are given by a Poisson point process with intensity measure $\tilde{a}(\chi_t)dt \times Q_Y(d\psi)$. Now let $(V_t)_{t \in [0, \infty)}$ be the total mass process of the associated tree of excursions with initial island measure that equals the distribution of Y in (4.166) and excursion measure Q_Y .

4.1.7.2 Condition for extinction of an invading defense allele

Proposition 4.1.17. *Assume the setting of Section 4.1.7.1. Let $x \in (0, 1]$ and assume $Y_0 = x = V_0$. Then the total mass process dies out (i.e., converges in probability to zero as $t \rightarrow \infty$) if and only if*

$$\alpha \geq \beta. \quad (4.167)$$

Proof. Define the functions $s: [0, 1] \rightarrow [0, \infty)$ and $S: [0, 1] \rightarrow [0, \infty)$ by $[0, 1] \ni z \mapsto s(z) := \exp\left(-\int_0^z \frac{-\frac{\kappa}{a}x(a-x) - \alpha x(1-x)}{\frac{1}{2}\beta(a-x)x(1-x)} dx\right)$ and $[0, 1] \ni y \mapsto S(y) := \int_0^y s(z) dz$. Note that for all $z \in [0, 1]$ it holds that

$$s(z) = \exp\left(\int_0^z \frac{2\kappa}{a\beta} \frac{1}{1-x} + \frac{2\alpha}{\beta} \frac{1}{a-x} dx\right) = (1-z)^{\frac{-2\kappa}{a\beta}} \left(\frac{a-z}{a}\right)^{\frac{-2\alpha}{\beta}} \quad (4.168)$$

and

$$S(z) = \int_0^z s(x) dx \leq zs(z). \quad (4.169)$$

The result will be shown using Theorem 5 from Hutzenthaler (2009). First, verify that the assumptions of the aforementioned theorem are satisfied. Using (4.169) yields

$$\int_0^{\frac{1}{2}} S(y) \frac{2}{\beta(a-y)y(1-y)s(y)} dy \leq \int_0^{\frac{1}{2}} \frac{2}{\beta(a-y)(1-y)} dy \leq \frac{1}{2} \frac{2}{\beta(a-\frac{1}{2})(1-\frac{1}{2})} < \infty. \quad (4.170)$$

Furthermore, it holds that

$$\begin{aligned}
 \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{\frac{1}{2}} \frac{-\frac{\kappa}{a}(a-y)y - \alpha y(1-y)}{\frac{1}{2}\beta(a-y)y(1-y)} dy &= \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{\frac{1}{2}} \frac{-2\kappa}{a\beta(1-y)} - \frac{2\alpha}{\beta(a-y)} dy \\
 &= \lim_{\varepsilon \rightarrow 0} \left(\frac{2\kappa}{a\beta} (\ln(1 - \frac{1}{2}) - \ln(1 - \varepsilon)) + \frac{2\alpha}{\beta} (\ln(a - \frac{1}{2}) - \ln(a - \varepsilon)) \right) \\
 &= \frac{2\kappa}{a\beta} \ln(1 - \frac{1}{2}) + \frac{2\alpha}{\beta} (\ln(a - \frac{1}{2}) - \ln(a)) \in (-\infty, \infty).
 \end{aligned} \tag{4.171}$$

From (4.168) as well as the fact that $\frac{2\kappa}{a\beta} > 0$ it follows that

$$\begin{aligned}
 \int_{\frac{1}{2}}^1 \frac{\tilde{a}(y)}{\frac{1}{2}\beta(a-y)y(1-y)s(y)} dy &= \int_{\frac{1}{2}}^1 \frac{\kappa a \frac{y}{a-y}}{\frac{1}{2}\beta(a-y)y(1-y)} (1-y)^{\frac{2\kappa}{a\beta}} \left(\frac{a-y}{a} \right)^{\frac{2\alpha}{\beta}} dy \\
 &= 2 \frac{\kappa a}{\beta} a^{-\frac{2\alpha}{\beta}} \int_{\frac{1}{2}}^1 (1-y)^{\frac{2\kappa}{a\beta}-1} (a-y)^{\frac{2\alpha}{\beta}-2} dy \\
 &\leq 2 \frac{\kappa a}{\beta} a^{-\frac{2\alpha}{\beta}} \left((a - \frac{1}{2})^{\frac{2\alpha}{\beta}-2} + (a-1)^{\frac{2\alpha}{\beta}-2} \right) \int_{\frac{1}{2}}^1 (1-y)^{\frac{2\kappa}{a\beta}-1} dy < \infty.
 \end{aligned} \tag{4.172}$$

It follows from (4.170), (4.171), and (4.172) together with a straightforward adaptation of Lemmas 9.6, 9.9, and 9.10 in Hutzenthaler (2009) to the state space $[0, 1]$ that the assumptions of Theorem 5 in Hutzenthaler (2009) are satisfied. Applying the aforementioned theorem shows that the total mass process dies out if and only if

$$\int_0^\infty \int_0^\infty \tilde{a}(\chi_t) dt Q_Y(d\chi) \leq 1. \tag{4.173}$$

Moreover, a straightforward adaptation of Lemma 9.8 in Hutzenthaler (2009) to the state space $[0, 1]$ together with (4.168) shows that

$$\int_0^\infty \int_0^\infty \tilde{a}(\chi_t) dt Q_Y(d\chi) = \int_0^1 \frac{\kappa a \frac{y}{a-y}}{\frac{1}{2}\beta(a-y)y(1-y)} (1-y)^{\frac{2\kappa}{a\beta}} \left(\frac{a-y}{a} \right)^{\frac{2\alpha}{\beta}} dy. \tag{4.174}$$

Observe that $\frac{2\kappa}{a\beta} \int_0^1 (1-y)^{\frac{2\kappa}{a\beta}-1} dy = 1$. Combining this with (4.173) and (4.174) it follows that the total mass process dies out if and only if

$$\begin{aligned}
 0 &\geq \int_0^1 \frac{\kappa a \frac{y}{a-y}}{\frac{1}{2}\beta(a-y)y(1-y)} (1-y)^{\frac{2\kappa}{a\beta}} \left(\frac{a-y}{a} \right)^{\frac{2\alpha}{\beta}} dy - 1 \\
 &= \frac{2\kappa}{a\beta} \int_0^1 (1-y)^{\frac{2\kappa}{a\beta}-1} \left(\frac{a-y}{a} \right)^{\frac{2\alpha}{\beta}-2} dy - 1 \\
 &= \frac{2\kappa}{a\beta} \int_0^1 (1-y)^{\frac{2\kappa}{a\beta}-1} \left(\left(\frac{a-y}{a} \right)^{\frac{2\alpha}{\beta}-2} - 1 \right) dy.
 \end{aligned} \tag{4.175}$$

Consequently, the total mass process dies out if and only if $\alpha \geq \beta$. This completes the proof of Proposition 4.1.17. \square

Chapter 5

Discussion

Factors influencing the success of altruistic defense traits

Understanding the evolution of altruistic behavior is an important topic in evolutionary biology. The theories of kin selection, group selection, game theory, or the Price equation can provide predictions of the success of altruism in a general setting. However, in the context of altruistic defense behavior against a parasite or predator, these methods may be difficult to apply. The interactions between host and parasite (or prey and predator) lead to complicated dynamics, especially in a structured setting, that can make it difficult to entangle the effects of altruistic defense behavior, and thus, to predict the success of altruists. The goal of the mathematical analysis was to understand the dynamics of altruists in a simple individual-based model for interactions of host and parasite or, equivalently, prey and predator. For the diffusion approximation of this model, it was possible to derive an expression for the frequency of altruists, as given by equation (3.2). This process was derived in Theorem 3.1.2 in the limit of infinitely large populations under a complete separation of ecological and evolutionary time scales. After taking the limits of infinitely many demes and an infinite time horizon, a simple condition determining the success of altruists was obtained. This condition allows to predict whether altruists will go to fixation or extinction depending on the ecological model parameters. Theorem 3.1.3 shows that in the limit of large populations structured in many demes, the altruist frequency will go to zero, if $\alpha > \beta$ and it will go to one, if $\alpha < \beta$. In the critical case of $\alpha = \beta$, there is coexistence between altruists and cheaters and the equilibrium distribution for the altruist frequency is given by equation (3.5).

The term β is defined as

$$\beta = \rho_{\frac{\delta}{\delta\nu + \lambda\gamma}} \beta_H, \quad (5.1)$$

which can be interpreted as the benefit of defense, scaled with the relative parasite pressure.

The parameter ρ denotes the decrease in parasite growth rate per altruist, and thus it is a measure of the effect of the defense allele. Furthermore, δ is the host death rate per parasite, ν denotes the parasite death rate, λ defines the host birth rate, γ is the competition rate among parasites, and β_H describes the randomness in host reproduction. Thus, the factor $\frac{\delta}{\delta\nu+\lambda\gamma}$ denotes the relative strength of parasites compared to hosts, i.e., the larger this factor is, the bigger the potential harm that hosts suffer from parasites. This represents the selective pressure on the host population to evolve defense mechanisms. If the pressure is strong, then the potential benefit from defending is high. In addition, randomness in host reproduction is beneficial to the evolution of the altruistic defense allele.

Débarre et al. (2012) investigate the evolution of altruistic defense traits in a host-parasite system with population structure. They conclude that the altruistic trait of suicide upon infection with an additional fecundity cost for the ability to detect the parasite can evolve in a structured population if the parasite is sufficiently harmful to the host population. This is in line with the definition of β in the present model that captures the parasite-induced selection pressure on hosts as well as the necessity of κ_H being positive, which ensures the population structure in the present model.

Notably, there are several parameters— κ_H , κ_P , β_P , η , and K —that do not enter the condition for fixation or extinction of the defense allele. At first it appears surprising that the amount of host migration is not relevant for the success of the altruistic allele, as long as it is positive. Berngruber et al. (2013) conclude that intermediate levels of mixing are necessary for suicide defense against infection to evolve in bacteria. Other studies also find that too much gene flow can hinder the success of altruism (Maynard Smith, 1964; Wade, 1978; Slatkin and Wade, 1978). In the present model, the parameter scaling ensures that the migration rate is small compared to the population size in the limit of large populations and thus, the amount of migration is in the appropriate range, irrespective of the value of $\kappa_H \in (0, \infty)$. Parasites mediate the altruistic effects and for this role the strength of their fluctuations or migration rates do not seem to be important, as long as they are in the required range that is achieved by the population size scaling in the limit of $N \rightarrow \infty$.

These results build on the assumption that the defense allele is already present in positive frequency in all demes and the evolutionary stability of altruism is investigated. This does not cover the case of a newly arisen defense allele that tries to invade the host population. Invasion is treated in Proposition 4.1.17. There, the defense allele is assumed to exist in positive frequency only in a single deme and the condition for a positive survival probability in a population structured in infinitely many demes is derived. Notably, a similar condition as for the case of evolutionary stability is retrieved and the survival probability is positive if and only if $\beta > \alpha$. A positive probability of survival does not ensure that the newly arisen

allele does, in fact, survive. So even if the condition $\beta > \alpha$ is satisfied, there may still be a chance that the altruistic defense allele goes extinct before reaching a positive frequency in the population. Nevertheless, the result states that if $\beta \leq \alpha$, then the defense allele will never be able to reach and sustain a positive frequency in a population structured in infinitely many demes. In a setting of finite population sizes, the probability of the allele frequency crossing some small threshold value is always positive, because of the pronounced impact of random fluctuations at low frequencies. Thus, the result is not as substantial in the finite setting, as it is in the case of large populations in many demes. Nevertheless, it gives the qualitative insight that it is possible for the altruistic allele to invade a population of cheaters.

The starting point of Hutzenthaler et al. (2015) is the diffusion model given in Section 4.1.3 and it is only mentioned that this diffusion model can be obtained from an individual-based model via a diffusion approximation. Section 2.1.1 and Section 4.1.2 close this gap by explicitly defining the individual-based model and providing the diffusion approximation.

Theorem 3.1.1 extends well-known results on Lotka-Volterra equations. To the best of my knowledge, this result is new even in the panmictic setting. If in the setting of Section 4.1.3 it is assumed for some $N \in \mathbb{N}$ that $\kappa_H^N = \kappa_P^N = \iota_H^N = \iota_P^N = \alpha^N = \beta_H^N = \beta_P^N = 0$ and that $A_0^N(i) = 0, i \in \mathcal{D}$, then for all $i \in \mathcal{D}$ it holds that $A^N(i) \equiv 0$ and $(H^N(i), P^N(i))$ satisfy classical Lotka-Volterra equations (Lotka, 1920; Volterra, 1926). Thus, if $K\eta > \nu$, then these equations have the nontrivial equilibrium $(\frac{K\delta\nu}{\lambda\gamma + K\delta\eta}, \frac{\lambda K\eta - \lambda\nu}{\lambda\gamma + \delta K\eta}) \in (0, \infty)^2$. Due to the assumptions of weak migration, weak immigration, weak selection, and low genetic drift, in the limit of large population sizes, the processes $(H^N(i), P^N(i)), i \in \mathcal{D}$, can be approximated by these classical equilibrium values, if η is replaced by $\eta - \rho F^N(i)$. Thus, the result extends the classical case to two host types—altruistic defenders and cheaters—with genetic drift in the spatial setting. These findings are formalized in Theorem 3.1.1.

Effect of finite population sizes

There are several approximations necessary in order to obtain the condition for the fixation or extinction of altruists. These approximations include observing infinitely large populations living in infinitely many demes for an infinite amount of time (cf. Figure 3.1). All of these conditions will obviously be violated when investigating any natural population. Furthermore, the assumption of a complete separation of ecological and evolutionary time scales implies that changes in altruist frequency immediately affect the host-parasite dynamics. Hence, there is no time lag between an act of altruistic defense and its benefit to the local host population. Still—without these idealizations being fully realistic—it is possible that the condition provides a good approximation for sufficiently large finite populations liv-

ing in finitely many demes that are observed for a limited amount of time. This was tested with the simulation models described in Section 2.2. In the case of small population sizes of the simulation model described in Section 2.2.1.1, the theoretical prediction is qualitatively correct. Namely, there is either extinction or fixation of altruists, with a narrow transition range between the two outcomes. However, the value for which this transition occurs does not coincide with the one predicted by the condition derived in Section 3.1. This is due to the fact that there is a large amount of inherent randomness when dealing with finite population sizes. The effect can clearly be observed in *simulation series* α . Even without altruism providing any “benefit”—i.e., $b = 0$, with $b = \rho\delta/\delta\nu + \lambda\gamma = \beta/\beta_H$ —the selective disadvantage of altruism needs to be relatively high in order for altruist extinction to be a stable result across simulations (see Figure 3.2). In 3 of the 100 runs with $b = 0$ and $\alpha = 0.2$ altruistic defense went to fixation. This effect was considerably more pronounced for positive (but relatively low) values of b . In general, in the finite setting, altruists are more highly favored than predicted by the asymptotic result. The effect of stochasticity in finite populations is captured by *simulation series* β_H , which evaluates the process for different values of the amount of randomness in host reproduction, β_H (see Figure 3.3). Even for $\beta_H = 0$, altruism can be favored when b is sufficiently large. In the diffusion approximation, all randomness based on the finiteness of the individual-based model disappears. Thereby, it relies on a positive value of β_H , to achieve differences between demes and β_H directly enters the condition for fixation or extinction of altruists. Thus, altruistic defense would always go extinct in the mathematical model, if $\beta_H = 0$. Apart from that difference, the qualitative influence of β_H in the simulations behaves according to the theoretical prediction, i.e., larger values of β_H lead to a stronger benefit of altruists (see Figure 3.3).

The asymptotic result predicts that there is no influence of the host migration rate κ_H on the evolution of altruistic defense (as long as $\kappa_H > 0$). In the context of small population sizes, migration rates may play a more important role than for the setting of large populations. Ecological and evolutionary time scales are not fully separated in the finite setting and thus, a large value of κ_H may mean that migration events interfere with host-parasite interactions. On the other hand, a small value may result in very little mixing of the population and thereby diminish a possible benefit of altruists. However, no significant effect of the host migration rate could be detected in *simulation series* κ_H (see Figure 3.4), even though values of κ_H range from 0.01 to 10, hence, spanning a 1000-fold difference. Note that this does not necessarily mean that there is no effect of κ_H . A significant influence might be found if more simulations were performed. Nevertheless, the amount of gene flow under which the evolution of altruism is favored may not be as limited in the present study as suggested by other approaches (Maynard Smith, 1964; Wade, 1978).

The positive effect of random differences in population sizes between demes on the success of altruists is also visible in *simulation series a_f* . Random fluctuations in the carrying capacities of the host and parasite populations were added in different amounts. Even though changes in carrying capacities affect both altruists and cheaters, they may be beneficial for the altruistic defense allele. This is due to the fact that, similar to increasing β_H , the random fluctuations in carrying capacities amplify differences between demes. These increased differences may lead to a more pronounced Yule-Simpson effect. Indeed, larger fluctuations lead to a benefit for the altruistic defense allele (see Figure 3.5). For the investigated parameter ranges, this effect is not as large as that of changing β_H , though. But note that stronger or more frequent fluctuations may provide even larger benefits for the defense allele.

In *simulation series β_H* , *simulation series κ_H* , and *simulation series a_f* , fixation or extinction of altruists is very stable across different simulation runs with the same parameter settings, except for the narrow transition ranges between the two outcomes. Thus, depending on the parameter setting, there is a clear advantage of either altruists or cheaters that can overcome the randomness that may lead to fixation of a disfavored type in a finite setting. This result is in line with the theoretical prediction. However, the exact value where the transition from an advantage of cheaters to an advantage of altruists occurs differs. This may be caused by the increased effect of randomness in the finite setting.

Effect of population structure

Simulations assuming infinite population sizes were performed in order to compare the success of altruists in finite and infinite settings and to investigate effects of population structure. The process X , which describes the relative frequencies of altruists in the limit of large population sizes under a complete separation of time scales according to equation (3.2) (see Theorem 3.1.2), was simulated.

The discretization of the process X makes it necessary to introduce a cutoff value ε , such that in each deme, values of X below ε or above $1 - \varepsilon$ are set to zero or one, respectively (see Box 2.5). In *simulation series ε* the appropriate choice of ε was determined to be 10^{-7} , which was the only value among the tested ones that showed a root mean squared error of zero for all simulated values of β (see Figure 3.6 for the outcome of all runs). For simulations with $\varepsilon = 0$, extinction or fixation of altruists never occurred within the given time frame of $T = 2000$, while for $\varepsilon = 10^{-8}$, it only occurred for the smallest value of β , but never for any other value. Thus, for $\varepsilon = 0$ and $\varepsilon = 10^{-8}$, the results indicate a quasi-equilibrium of altruists and cheaters, with a large altruist frequency for $\beta > \alpha$ and a low altruist frequency for $\beta < \alpha$. On the other hand, for $\varepsilon = 10^{-6}$, the final altruist frequency was zero or one in

97 out of 100 runs. In these cases, the outcome always matched the prediction. However, for three simulation runs with $\beta > \alpha$, a relatively low altruist frequency was observed at the end of the simulation run. Note that if the process runs for a very long time, then the outcome may still coincide with the prediction. Only for $\varepsilon = 10^{-7}$, all simulations agreed with the theoretical predictions and fixation or extinction of altruists was reached in the given time frame.

The simulations with uniform migration and $D = 500$ fit the theoretical prediction extremely well (see Figure 3.9). In the mathematical analysis, uniform migration is assumed, because it is necessary for the mean field approximation (see Section 4.1.5). Nevertheless, for the other migration schemes investigated, the prediction was also closely met, when the number of demes was large enough (i.e., 500, 484, or 511). For these other migration schemes, fixation or extinction of altruists was often not reached within the given time frame when β was very close to α . This may be explained by the fact that under uniform migration, population mixing is strongest and any given deme can best be reached by individuals.

In several simulation series, fixation or extinction of altruists was also observed for $\beta = \alpha$, contrary to the theoretical predictions. Because of the finite number of demes, there is always a positive probability of one type going extinct. This probability gets smaller with an increasing number of demes, as the simulations show (see e.g., Figure 3.9).

For intermediate values of D (i.e., 50, 49, or 63), fixation or extinction contradicting the prediction could be seen in several runs. This never occurred for the largest values of D (500, 484, or 511), though. All runs with a final altruist frequency of zero or one were in accordance with the prediction. If the simulations had been kept running longer, then a fixation or extinction contradicting the prediction might still have happened. However, the larger the number of demes, the more unlikely it gets to observe such a contradiction to the prediction.

Across all simulation series it is apparent that less than ten demes are not enough for altruists to be strongly favored. For all four migration patterns that were investigated, the mean final altruist frequency over the 100 simulation runs never exceeded 0.8. This is also reflected in the relatively high root mean squared errors for the lowest values of D in the four simulation series.

For uniform migration and nearest neighbor migration on a circle, extinction of altruists was stable across the 100 simulations, if β is sufficiently small. This was not the case for nearest neighbor migration in two dimensions and migration along the edges of a binary tree. Indeed, for *simulation series 2D* and *simulation series T*, the final altruist frequency stayed strictly within the boundaries of zero and one for all simulations runs across all parameter settings. Thus, there is an effect of the specific form of how the population is structured.

Migration between nearest neighbors in two dimensions and along edges of a binary tree appears to prolong the time of coexistence between altruists and cheaters, compared to the uniform migration and nearest neighbor migration on a circle.

Effect of stochasticity

An altruistic act of defense leads to a deme-wide reduction of parasite pressure. Consequently, the equilibrium values of population sizes of hosts and parasites in the deme change, and the effective host carrying capacity increases, according to the function h_∞ (see equation 4.13). The larger a deme is, the more migrants it sends to other demes. Thus, the increase in carrying capacity becomes a benefit for altruists, if there are smaller demes, with a high frequency of cheaters. Then, even though being disfavored within each deme, altruists can obtain an advantage when looking at the whole population. This reasoning is known as the Yule-Simpson effect or as Simpson's paradox (Yule, 1903; Simpson, 1951; Blyth, 1972). In the result from the mathematical analysis, stochastic effects are captured by the factor β_H in the “benefit” term $\beta = \rho \frac{\delta}{\delta\nu + \lambda\gamma} \beta_H$. Hence, randomness in host reproduction—which leads to differences between demes, and thus makes the Yule-Simpson effect possible—directly affects the evolution of the altruistic defense allele. Findings from the simulation studies confirm this view. In the setting of finite population sizes, randomness in host reproduction is increased. Consequently, between-deme differences are enhanced, which results in altruism being more strongly favored than predicted by the mathematical analysis. This effect of randomness is further shown by varying the value of β_H in the simulations (see Figure 3.3) as well as adding random fluctuations to the carrying capacities of host and parasite (see Figure 3.5).

Overall, simulation results of the process X approximate the theoretical prediction well, if the number of demes is sufficiently large. Due to the assumption of infinite population sizes and a complete separation of time scales, random effects are reduced to those introduced by β and by the finite number of demes. Hence, the simulations of X fit the prediction better than the simulations of the individual-based model. Nevertheless, the theoretical prediction is insightful for the finite setting, in which the transition from altruist fixation to extinction is shifted in favor of altruists due to a larger impact of stochasticity. This is in line with the original result that predicts a strong influence of randomness, captured by the factor β_H in the term $\beta = \rho \frac{\delta}{\delta\nu + \lambda\gamma} \beta_H$.

The role of stochasticity in evolutionary processes has long been recognized (see Lenormand et al., 2009 for a review). Finite population sizes are an important source of stochasticity and lead to so-called “genetic drift”. In a finite population with random reproduction,

some individuals will have disproportionately large success in producing offspring, while others leave few or no descendants. The larger the population gets, the smaller the effect of this random sampling will be on the development of allele frequencies over time. The impact of genetic drift is typically modeled to be inversely proportional to the population size, such that it can be neglected in a sufficiently large population. This is why many models assume large population sizes and then proceed to describe the evolution of the population in a deterministic fashion. In the context of host-parasite interactions, a large amount of studies models the numbers of susceptible and infected individuals of a host population by deterministic differential equations (Boots and Haraguchi, 1999; Boots and Bowers, 2004; Miller et al., 2007; Best et al., 2009; Berngruber et al., 2013). Typically, this is an adequate approach for understanding host-parasite dynamics, but there may be phenomena that require the consideration of randomness to obtain a suitable model. One example is that of rare events, as shown by Handel et al. (2009). They investigate the creation of antiviral resistance and show that a stochastic model produces different results than the deterministic approximation. Even though large populations are considered in the present model, a deterministic approach would not yield the correct results. Ignoring all stochasticity for large populations neglects the fact that very small random effects can add up over time and become noticeable. Hence, viewing the same process on a larger time scale reveals the influence of very small evolutionary forces. This is captured in Section 4.1.4, where the forces of genetic drift, migration, and selection are negligible on the original ecological time scale, but once the processes are sped up and thus investigated on the much larger evolutionary time scale, these forces prove to be relevant. In Theorem 3.1.2 the convergence of the altruist frequency on the evolutionary time scale in the limit of large populations is shown and the limit process captures effects of genetic drift, migration, and selection. The result demonstrates the importance of random effects. If all stochasticity in host reproduction was excluded from the diffusion model 4.9—i.e., $\beta_H = 0$ —then the altruist would always go extinct. For the evolution of altruists in the present model, it is necessary that there are random fluctuations around the expected values. These fluctuations can produce differences between demes that lead to an advantage for demes with a larger altruist frequency. Note that random disturbance of average effects needs to be an ongoing process, and it is not sufficient to simply start with random frequencies. The maintenance of between-deme variation by the random forces of genetic drift, fluctuating environments, and mutation and its role in the evolution of altruism is discussed in various studies (Lande, 1976; Slatkin and Wade, 1978; Uyenoyama, 1979, discussed below).

Parvinen et al. (2003) demonstrate differences between a stochastic model and its deterministic approximation when investigating the evolution of dispersal rates. When local

population sizes fluctuate around their carrying capacity, it can be beneficial for an individual to move to a deme with more available resources. This effect is not captured by deterministic models. Randomness does not simply add noise to average results, but can qualitatively change the outcome of a model. Nevertheless, the approach of a deterministic approximation leads to important insights into evolutionary dynamics and, hence, is a standard approach used in a large number of studies (see Parvinen et al., 2003 for an overview).

Relation to kin selection, group selection, and game theory

As discussed above, random differences between demes, as obtained by genetic drift in the present model, can invoke the Yule-Simpson effect (Yule, 1903; Simpson, 1951; Blyth, 1972). Its connection to group selection theory has already been discussed by Sober and Wilson (1998) and Gardner (2015). There are, in fact, two different concepts of group (or multilevel) selection, as discussed by Okasha (2006) and Gardner (2015). When studying a population of individuals living in groups, the different notions arise from the question what constitutes the focal level. If the goal is to understand how frequencies of different types of individuals in the population change, then the focus is on the individual. This concept is known as multilevel selection 1 or MLS-1. On the other hand, if the analysis aims at understanding how frequencies of different types of groups change, then the focus is on the group. This notion is referred to as multilevel selection 2 or MLS-2 (Okasha, 2006; Gardner, 2015). Thus, the present analysis of altruistic defenders is in line with the concept of MLS-1. The question is how frequencies of different types of individuals change and there is no plausible notion of reproducing groups, which would be necessary for MLS-2 to be invoked.

The present result extends earlier analytical approaches considering the evolution of altruism under group selection to the case of individual-based migration in the context of defense against a parasite. Uyenoyama (1979) investigates the evolution of altruism in an island model, assuming a stage of random mixing of the whole population followed by recolonization of islands in each generation. This model goes back to Levene (1953) and has also been utilized by Gillespie (1974). The simplification of assuming a stage of random mixing allows to immediately obtain a one-dimensional diffusion approximation. This assumption is removed in the present analysis, which drastically increases the complexity of deriving a one-dimensional diffusion for the frequency of altruists (Sections 4.1.2 to 4.1.5). In line with MLS-1, Uyenoyama (1979) concludes that for group selection to act, it is not necessary to invoke a mechanism for complete extinction of whole groups, but instead, it is enough if there is maintained variation between the groups, e.g., due to genetic drift or fluctuating environments. Also note that Lande (1976) explains how mutation can help in

maintaining genetic variability in a polygenic character. In the present model, variation between groups is maintained solely via genetic drift, which is a sufficient force even in large populations, when the processes are analyzed on a large enough time scale. Thus, if some of the other mechanisms to increase and maintain genetic variability (as discussed by Lande, 1976; Uyenoyama, 1979) are introduced in the model, this might in fact provide an even larger benefit for altruism.

In the model of Slatkin and Wade (1978) whole demes may go extinct with a certain probability in each generation. Selection acts on the level of individuals within the demes and there is gene flow between all demes. These two forces act to decrease between-deme variation, while mutation (Lande, 1976) and genetic drift increase it. Two different approaches of how empty sites may be recolonized are investigated. The founding population can either be sampled randomly from the whole population or it stems completely from a single deme, chosen at random. For the character at interest, the equilibrium values for its mean and between-deme variance are computed. It is concluded that recolonization from single demes increases between-deme variance. Furthermore, the impact of genetic drift is largest when the effective population size per deme is small and the probability of extinction per generation is low, as this increases the time that drift has to show an effect. However, variation between demes can also be increased by a strong founder effect, caused by small founding population sizes and high extinction probabilities, especially when vacant sites are colonized from single demes. In the present study, the assumptions of Slatkin and Wade (1978) (small population sizes per deme, extinction of demes, mutation, migration in groups) are not satisfied, yet, genetic drift produces enough variation between demes in order for deme-level selection to counteract the cost of altruism, when looking at a large enough time scale (as discussed in the section on the effect of stochasticity). As Wade (1978) notes, assuming that all demes contribute an equal amount of migrants independent of their size, can reduce the effects of selection on the group level. This is in agreement with the present study, for which allowing larger demes to send out more migrants is the key part in the success of altruists, because demes with a high altruist frequency tend to be larger. Maynard Smith (1964) concludes that migration is detrimental to the strength of group selection, because it reduces variance between demes. As Wade (1978) points out, this reasoning is based on the large amount of individual selection assumed in the haystack model of Maynard Smith (1964), where a single non-altruistic individual takes over a local population of altruists. Maynard Smith (1964) notes that this model is very artificial and due to its extreme assumptions it may not provide convincing evidence for the evolution of altruism by means of group selection. However, later studies find that group selection is relevant also under less restricting conditions (Wade, 1978; Slatkin and Wade, 1978; Uyenoyama, 1979). This view is also

supported by the results of *simulation series* κ_H , which do not show a significant effect of the host migration rate (Figure 3.4). Thus, even moderate levels of mixing do not seem to hinder the success of altruism in the present approach.

The same effect can also be interpreted according to the theory of kin selection. The relatedness term in Hamilton's rule captures genetic similarity between the focal individual and its interaction partners, regardless of how that similarity arises. Birch and Okasha (2015) state that the most common source of such genetic correlation between partners is kinship, but it may also stem from other mechanisms, such as greenbeard effects (Dawkins, 1976; Gardner and West, 2010). Relatedness is strictly defined as the coefficient of a least squares linear regression analysis, in which the genetic value of the focal individual's partner is predicted by the genetic value of the focal individual's character of interest (Hamilton, 1963, 1964a; Gardner et al., 2011). Denoting by g the actor's genetic value for the character of interest and by g' that of the actor's social partner, relatedness can be expressed as $\text{Cov}(g, g') / \text{Cov}(g, g)$, where covariances are taken over the whole population (Gardner et al., 2011). If fitness effects are mediated by phenotypes, then this formulation can be adapted to $\text{Cov}(g, p') / \text{Cov}(g, p)$, where p and p' denote phenotypic values of actor and recipient, respectively (Orlove and Wood, 1978; Queller, 1992). Informally, relatedness can be seen as an indicator of how likely it is that a focal gene of the actor is found in the recipient, in relation to the population average, which is given by the average gene frequency over the whole population (Bourke, 2011). Thus, a value of zero means that the gene frequency among recipients equals the population average, while a value of one means that all recipients share the focal gene. Intermediate values are derived accordingly, assuming a large population (Grafen, 1985). Furthermore, the relatedness term can also take negative values, if it is less likely that the focal individual shares the gene in question with the interaction partner, than with a randomly sampled individual. A geometrical representation of this concept of relatedness is given by Grafen (1985). In the simplified model presented here, there is no distinction between individuals of the same type and all altruists are basically clones. Furthermore, effects of an increased carrying capacity influence all individuals in the deme equally. Thereby, a high altruist frequency in any given deme corresponds to a high relatedness of an altruist to other individuals in that deme that all gain the same benefit from the defense behavior of the altruist. Differences between demes, which rely on varying altruist frequencies and changes in carrying capacity, can thus be analogously explained by relatedness and benefits of altruistic defense.

However, high amounts of relatedness in a structured population can also mean that relatives compete for limited space or resources. This competition has been shown to potentially cancel the beneficial effects of cooperation and thus prevent the evolution of altruism (Wil-

son et al., 1992; Taylor, 1992). However, Alizon and Taylor (2008) show how this effect of competition can be reduced by allowing for empty sites in local populations and varying dispersal rates depending on the size of a deme. Thereby, migration increases with increasing density in a deme and competition can be reduced. In fact, Alizon and Taylor (2008) show that Hamilton's rule holds in their model when an infinite number of demes is assumed. Van Baalen and Rand (1998) investigate a lattice model, in which each node can harbor an altruist or a cheater or it can be empty. They specify rates at which each pair of connected nodes change their states due to a birth, death, or migration event. Using the approach of "correlation dynamics" (Durrett and Levin, 1994; Durrett, 1995), this model can be analyzed using differential equations describing the expected rates of change of the proportion of sites (or pairs of sites) in different states. By allowing for empty sites, clusters of altruists can grow bigger than clusters of non-altruists and thus, altruism can evolve. Furthermore, Van Baalen and Rand (1998) obtain the validity of Hamilton's rule for their model. Platt and Bever (2009) review the mechanisms that reduce effects of kin competition and allow for a spread of altruism in structured populations. Apart from allowing for empty sites, an important concept is that of population elasticity. By allowing local carrying capacities to increase under a high density of cooperators, the effect of kin competition can be reduced. Other factors include kin recognition (greenbeards) and kin-structured dispersal to increase relatedness without increasing kin competition. Moreover, if dispersal occurs immediately after cooperation, then benefits may be shared between cooperators, while kin competition is reduced (Platt and Bever, 2009). Lastly, in a group-structured population, increased productivity of cooperative groups may favor altruism via between-group selection (Platt and Bever, 2009). In the present model, kin competition is reduced due to population elasticity. Here, a high altruist frequency reduces parasite pressure and thereby increases the potential local host population size.

In the present approach, it is not necessary to decompose effects into cost, benefit, and relatedness terms or into components acting within and between groups. Instead, the altruist frequency is described in a diffusion process and the tools from diffusion theory are used to analyze this process and determine its longterm behavior. As Berngruber et al. (2013) point out, the success of an altruistic defense trait depends on parasite presence. Thus, benefits of altruistic acts are not constant in the scenario of defense. These fluctuations of effect strengths as well as changing values of relatedness do not need to be considered using the present result. Notably, the current state of host and parasite populations does not enter in the condition for the success of altruism. Instead, the advantage or disadvantage of the defense trait remains fixed throughout its evolution, if all demes start with the same distribution of altruist frequency.

Similar to game theoretic approaches such as the Prisoner's dilemma (Axelrod and Hamilton, 1981), cheating appears to be the favorable strategy in the present analysis. Independent of the types of other individuals in a deme, a focal individual has an increased immediate reproductive success by cheating compared to cooperating. Nevertheless, the results show that, in fact, altruism will prevail under suitable conditions. Several explanations have been given for such seemingly contradicting results using the methods of game theory. Including spatial structure in evolutionary games can lead to the evolution of altruism (e.g., Ohtsuki et al., 2006). In the context of the Prisoner's dilemma, the effect of population structure on the success of cooperation has also been recognized (Fletcher and Zwick, 2007). This, of course, is reflected in the present analysis, which also relies on spatial structure and differences among demes to achieve an advantage of altruists. Furthermore, introducing repeated interactions has been shown to increase the long-term success of cooperators compared to cheaters when considering the Prisoner's dilemma (Trivers, 1971; Nowak, 2006). In the present modeling approach, this insight may be applied as well. There are not merely single interactions between individuals, but instead partners will stay together (and interact) for considerable amounts of time. This is achieved by a relatively weak migration process, such that individuals do not change demes too quickly. Thus, the present results can also be interpreted in the light of repeated interactions, for which cooperating is beneficial in the long run (under suitable conditions), while cheating provides only short-term benefits. Débarre et al. (2014) investigate a lattice model of fixed size with one individual at each site. The mechanisms for dispersal and social interactions (in the form of the Prisoner's dilemma and other games) are separated and the model allows for birth-death updating (i.e., first choose an individual to reproduce, then determine which individual gets replaced) and death-birth updating (i.e., first determine which individual dies, then choose an individual to place its offspring at the corresponding site). The choice between these updating mechanisms seems arbitrary, but it can influence the success of altruism (Ohtsuki et al., 2006; Taylor et al., 2007). In the approach of Débarre et al. (2014), social interactions can influence both steps of either updating scheme. They find that altruism is favored, if it affects the second step. In that case, social interactions range over two edges of the lattice and direct competition between relatives is reduced. Otherwise, the benefits are negated by kin competition effects and thus, spatial structure does not increase the chances of the evolution of altruism. This effect is based on the assumption of fixed population sizes. Hence, it is not visible in the present approach, where altruists increase the carrying capacity of the host population.

Negative feedback effects

The special case of altruism taking the form of defense against a parasite may lead to results that differ from the general case, because the fate of the altruist is linked to the fate of the parasite. If there is a high altruist frequency in a deme, this means that there are relatively few parasites. This could, in turn, reduce the benefit that altruists provide to the host population, and thereby lead to a reduction in altruist frequency. Consequently, there could be a negative feedback mechanism, which prevents the altruist frequency from growing beyond a certain level. This reasoning is in accordance with Sibly and Curnow (2011), who find that there can be coexistence of altruists and cheaters in a population if help is subject to diminishing returns. Thus, if there are already many cooperators in the population, then increasing the amount of cooperation even further will yield relatively little additional benefit. Interestingly, in the mathematical results, there is no such negative feedback mechanism and coexistence occurs only in the critical case $\beta = \alpha$. This can be explained by the function $h_\infty: [0, 1] \rightarrow (0, \infty)$, given by equation (4.13), which determines the equilibrium host population size in dependence of the altruist frequency. The second order derivative of h_∞ is positive, meaning that the additional benefit increases with the current value of the altruist frequency. Hence, altruism is not subject to diminishing returns in the present scenario. Consequently, the above-mentioned result from Sibly and Curnow (2011) does not apply and there will not be coexistence of altruists and cheaters. Thus, in the case that $\beta > \alpha$, altruists can reach fixation instead of fluctuating at an intermediate level. Nevertheless, an extreme case of negative feedback can be observed in *simulation series* β_H of the individual-based simulations. For the largest “benefit” value $b = 0.7$ (where $b = \beta/\beta_H$), the average final altruist frequency is lower than for $b = 0.6$ (see Figure 3.3). The strong defense of altruists in the case $b = 0.7$ actually leads to extinction of parasites in most cases (see Section 3.2.1). If this extinction of parasites occurs before the fixation of altruists, then altruism provides no further benefit and cheaters can take over the population. Thus, in this scenario, the benefit of defense can actually be detrimental to the evolution of altruism in the long run, due to the fact that parasites can go extinct in this finite system. This result is similar to the findings of Duncan et al. (2011), who show in *Paramecium Caudatum* that after removing contact to a parasite, costly resistance is no longer maintained and productivity of the host population slowly increases. Due to lack of mutation in the simulation model of Section 2.2.1.1 this is only apparent in runs where both cheaters and altruists are present after parasites go extinct. Notably, a negative feedback mechanism only occurs in this extreme scenario of *simulation series* β_H , while otherwise the situation resembles that of the mathematical analysis. It would be interesting to see if adapting the model such that effects of defense yield

diminishing returns, would allow for the possibility of coexistence and cheaters as predicted by Sibly and Curnow (2011).

Model assumptions

Several assumptions of the model suppose an ideal scenario, which will not occur in nature. Notably, the assumption that a single biallelic locus determines whether an individual behaves altruistically or not, is very strong. It might be more realistic, as well as possibly insightful, to extend the present modeling approaches to allow for a wider spectrum of defense strategies depending on several loci. Furthermore, individuals are assumed to be haploid to simplify the analysis. Thus, the dynamics of the altruistic allele may be altered when populations of different ploidy are considered. Another important assumption is the idealized structure of the populations. All demes are modeled to provide the same conditions for the populations and they are connected only by a slow migration process. An additional strong assumption is that of weak selection, weak migration, and weak genetic drift that is achieved by the parameter scaling in the limit of $N \rightarrow \infty$. This leads to the complete separation of time scales, as discussed above. Although such a split of ecological and evolutionary forces will not be as rigid in natural populations, it is a standard assumption in population genetics. Nevertheless, one has to be careful in interpreting the results and keep in mind that they rely on an idealized scenario. The effects of these assumptions can be seen in the differences between the simulation studies for finite and infinite population sizes (Sections 3.2.1 and 3.2.2).

Coevolutionary dynamics of host and parasite may lead to changes in strategies of virulence or resistance over time, especially in the context of spatially structured populations—as captured by the concept of a geographic mosaic of coevolution (Thompson, 1999; Best et al., 2009). Details of the coevolutionary process are not explicitly modeled in the present approach and thus, changes of strategies cannot be observed. Especially, as Best et al. (2009) show, parasite virulence can be highly dependent on evolutionary dynamics of the host, which may influence the outcome in the present approach. Nevertheless, important conclusions can be drawn from the current results. If there is adaptation of the parasite to the defense strategies of altruists, this may lead to complex dynamics in the spirit of a coevolutionary arms race (Dawkins and Krebs, 1979; Foitzik et al., 2003). Counter-adaptations of parasites to the defense strategies of altruists become relevant only after altruists have reached a high frequency. Since the present results predict that there is no coexistence of different host strategies, i.e., altruism and cheating, it is likely that fixation of altruism will occur shortly after parasites feel the pressure of evolving counter-adaptations. This may re-

sult in a step-wise arms race, allowing the present results to be applied successively. Still, an extension of the present modeling approach to explicitly allow for coevolution might be insightful.

Notably, no direct mechanism allowing altruists to preferentially cooperate with each other, such as kin recognition or greenbeards, is assumed in the present model. Furthermore, there is no explicit deme competition, as modeled in some classical group selection approaches such as the haystack model of Maynard Smith (1964). Instead, demes are connected only via (weak) migration of individuals. This is sufficient to compensate for the disadvantage that altruists suffer within any given deme, provided that $\beta > \alpha$. This relates to the findings of Slatkin and Wade (1978) and Uyenoyama (1979), discussed above.

The case of slave rebellion

A special case of altruistic defense in the form of ant slave rebellion was investigated with the simulation model described in Section 2.3.1. This first paragraph is based on the discussion by Metzler et al. (2015). Interactions of slavemaker ants and a host species were modeled on a grid that represented a forest area of $100\text{ m} \times 100\text{ m}$. First, data from field observations of *Protomognathus americanus* and *Temnothorax longispinosus* from New York State and West Virginia was combined with prior knowledge in an ABC approach to find plausible parameter combinations. Using these parameter settings, the possibility of a spread of the rebel allele, starting at a frequency of 1%, was investigated. The results show that a spread of rebellion is indeed possible, although only for few of the plausible parameter combinations. Pamminger et al. (2014) show that the evolution of slave rebellion could possibly be explained by kin selection on the level of individuals. Enslaved host workers could help their nearby relatives in unparasitized nests by reducing the raiding pressure via rebellion. If instead whole nests are viewed as “superorganisms” (Hölldobler and Wilson, 2009), then a kin selection argument may still apply. Close-by nests are related to each other (Pamminger et al., 2014) and thus, rebels in one nest can benefit their relatives in nests in close proximity. Thus, considering nests—or genes in queens and males of the nests—can provide a new perspective. Indeed, the results show that kin selection on the nest level can explain the evolution of rebellion, but only in a very small number of simulation runs and only if the cost of rebellion is extremely low (see Figure 3.14 for an example). The analysis of Pamminger et al. (2014) includes a simplified mathematical model. In the present approach, a lot of biological details have been included, such as modeling nest sizes and ages, or splitting of satellite nests. Due to this added complexity, the model can only be investigated with computer simulations. According to Pamminger et al. (2011), the cost of rebellion might

be between 5% and 10%. In the present simulations, a spread of rebellion could also be observed if the cost was 1% or higher. These cases did not involve a steady increase in rebel frequency over a longer period of time, though. Instead, the populations of host and parasite show a pattern of frequent extinctions and recolonizations. Thus, if there is a high rebel allele frequency among the founders during a recolonization, then rebels are wide-spread during the following years, which better protects the host population from slavemakers (examples are shown in Figure 3.13). Hence, if the rebel frequency is high, host nests can be more productive for a longer amount of time. This could lead to an advantage of rebels, if several patches are connected via migration to form a metapopulation. In fact, for several parameter combinations, this effect can be observed for simulations of a metapopulation of 20 connected patches. The rebel allele can get close to fixation and an analysis of Queller's group selection variant of the Price equation (Price, 1970; Queller, 1992) shows that the impact of the between-deme component is much stronger than that of the within-deme component (see Figure 3.17). This effect in the *Metapopulation simulations* could potentially be further increased by allowing not only the relative frequency of altruists among migrants to be determined by the average frequency across patches, but by also varying the amount of migrants depending on local population sizes. This may allow for a faster spread of altruists compared to the current approach. The results on the metapopulation can also be interpreted from a kin selection perspective (Hamilton, 1963, 1964a,b). In order for the rebel frequency in a patch to be high after a recolonization, it is necessary that the patch is populated by only a few queens and their offspring. Otherwise, if there are many queens among the founding population, then the rebel allele frequency among them will not be far from the average value in the metapopulation. Thus, relatedness between nests in the grid will be large shortly after a recolonization with a high rebel allele frequency occurred. Indeed, a positive correlation between the chance of the rebel allele to spread and the deme-wide host nest relatedness after recolonization was found. This leads to a large between-deme variation in allele frequencies, which is necessary for the Yule-Simpson effect to work (Yule, 1903; Simpson, 1951; Blyth, 1972). Thereby, the explanations derived from kin selection and group selection theory are tightly linked for the case of slave rebellion. The need for a quick growth during a recolonization of the grid explains the significant influence of the parameter *hostRelProductivity* that describes the growth rate of host nests. The significance of the parameter *Slavemaker-PoolDie* that represents the death rate of Slavemaker queens trying to conquer a host nest, can be explained similarly, as a high rebel frequency is only beneficial if slavemakers occur quickly. Otherwise, if slavemakers remain absent long enough, the cost of rebellion will result in a decline of rebel nests (see Figure 3.16).

In the simulation model for the evolution of slave rebellion, effects that multiple host

species (Brandt et al., 2005b) or multiple slavemaker species (Johnson and Herbers, 2006) may have, are ignored. These may impact the results, though, especially because the rebellion behavior relies upon differences in the cuticular hydrocarbons between host and parasite. If multiple host and parasite species are considered, differences could become more pronounced. This might lead to complicated dynamics, such as parasites shifting to a particular host species and the results could qualitatively be altered. Furthermore, it could be relevant whether the defense strategy is specific to a particular parasite or whether it targets multiple parasites at once (Frank, 2000). However, including multiple species of host and parasite would greatly increase the complexity of the simulation program and make it even more difficult to obtain meaningful insights into the evolution of the rebellion behavior.

An important assumption of the mathematical model is that extinctions of host and parasite do not occur in any given deme. However, frequent extinction and recolonization events are a key element in the simulations for the evolution of slave rebellion. Thus, these results cannot be interpreted in a meaningful way by the condition derived from the mathematical analysis. Similarly, in *simulation series* β_H of the individual-based model, extinction of parasites was observed in a large number of runs with the largest value of b , rendering the theoretical prediction meaningless. However, among the simulations of ants, there was a single parameter set that showed a stable equilibrium of slavemaker and host over the full 5000 years in ten out of ten simulations of the *Posterior simulations*. This parameter combination did not allow for the spread of rebellion, though, and was thus not investigated further. It may well be possible that combinations exist that allow for stable population sizes and a spread of rebellion. These might then be analyzed in the light of the asymptotic results. However, an extremely large number of simulations with the prior distributions would be necessary to further explore the huge parameter space. Despite the theoretical prediction not being directly applicable to the simulations of ant interactions, the general insights are still relevant. In particular, the effects of population structure, finite population sizes, and stochasticity, found in the analysis of an abstract altruistic defense trait, apply to the case of slave rebellion. The importance of population structure is shown by the impact of relatedness and the results of the *Metapopulation simulations*. Furthermore, the metapopulation study reveals that random differences between demes are necessary for the success of the altruist, which is in line with the insights from *simulation series* β_H and *simulation series* a_f . In finite populations, randomness plays a major role in the case of slave rebellion, where success of rebels relies on an over-representation of the rebel allele during a recolonization event—a phenomenon that could never be observed in a deterministic setting. This agrees with the conclusion of Slatkin and Wade (1978), who state that large extinction rates of local subpopulations together with small founding populations increase the between-deme variance in a structured population

and thus favor the evolution of altruism.

Conclusion

The present results show that an altruistic defense trait against a parasite or predator can prevail in a structured population under certain assumptions. Starting from an individual-based model of altruists, cheaters, and parasites, a condition for fixation or extinction of altruists was derived. This result provides insights into mechanisms that influence the success of altruistic defense traits. Randomness in host reproduction is necessary for the population structure to be relevant. Furthermore, the reduction of parasite pressure by altruistic defense does not lead to a negative feedback of the abundance of altruists. Instead, either altruists or cheaters are typically favored, and coexistence only occurs in the critical case of equality of the relevant parameters. The parasite acts as a mediator of the benefits provided by altruists to the host population, and migration rates or randomness in parasite reproduction are not influential. However, the relative strength of parasites is important and enters the “benefit” term β . The mathematical findings were confirmed by simulations and the impact of randomness was highlighted. For small populations or for populations that are divided in small subpopulations, stochasticity is more pronounced. This can be beneficial for the altruistic allele, which relies on randomness to create differences between demes that enable the Yule-Simpson effect. As shown in the simulation study, the evolution of altruism in the population structured in demes appears robust towards different intensities of gene flow. In particular, there is no detectable difference in the success of altruists for a large range of migration rates. However, the pattern of migration is shown to be highly relevant. Thus, in order to derive meaningful predictions for any natural population, the spatial structure has to be modeled appropriately. The simulations indicate that a sufficiently large number of demes is necessary for the results to apply. Otherwise, the effects of the population structure do not become relevant and the success of altruists is mainly determined by random effects. The simulations also indicate that the results are shifted when small populations are considered and when evolutionary and ecological time scales are not fully separated. This shows that although general insights may remain valid, results obtained from models with idealized assumptions need to be considered carefully when applied to natural populations. Applicability may be improved if model assumptions better reflect the details of the system that is investigated. In the simulation study on slavemaking ants, this is achieved by using field data to obtain reasonable parameter choices for the complex model. The insights from studying an abstract altruistic defense trait remain valid in the case of slave rebellion. In particular, it is shown that a costly allele for rebellion can be selectively favored under certain param-

ter combinations. Nevertheless, the mechanisms involved in the evolution of a hypothetical allele for slave rebellion can be understood in greater detail with the more complex model. The benefit of protection against slavemakers provided by rebellion is especially influential in a metapopulation scenario. In any case, random effects of small (founding) population sizes play an important role in the success of the defense allele.

Bibliography

- Abbot, P., Abe, J., Alcock, J., Alizon, S., Alpedrinha, J. A., Andersson, M., Andre, J.-B., van Baalen, M., Balloux, F., Balshine, S., et al. (2011). Inclusive fitness theory and eusociality. *Nature*, 471(7339):E1–E4.
- Achenbach, A. and Foitzik, S. (2009). First evidence for slave rebellion: enslaved ant workers systematically kill the brood of their social parasite *Protomognathus americanus*. *Evolution*, 63(4):1068–1075.
- Achenbach, A., Witte, V., and Foitzik, S. (2010). Brood exchange experiments and chemical analyses shed light on slave rebellion in ants. *Behav. Ecol.*, pages 948–956.
- Alizon, S. and Taylor, P. (2008). Empty sites can promote altruistic behavior. *Evolution*, 62(6):1335–1344.
- Alloway, T. M., Buschinger, A., Talbot, M., Stuart, R., and Thomas, C. (1982). Polygyny and polydomy in three North American species of the ant genus *Leptothorax* Mayr (Hymenoptera: Formicidae). *Psyche*, 89(3-4):249–274.
- Archetti, M. and Scheuring, I. (2012). Review: Game theory of public goods in one-shot social dilemmas without assortment. *J. Theor. Biol.*, 299:9–20.
- Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.
- Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4):2025–2035.
- Berngruber, T. W., Lion, S., and Gandon, S. (2013). Evolution of suicide as a defence strategy against pathogens in a spatially structured environment. *Ecol. Lett.*, 16(4):446–453.

BIBLIOGRAPHY

- Best, A., Webb, S., White, A., and Boots, M. (2011). Host resistance and coevolution in spatially structured populations. *Proc. R. Soc. B*, 278(1715):2216–2222.
- Best, A., White, A., and Boots, M. (2009). The implications of coevolutionary dynamics to host-parasite interactions. *Am. Nat.*, 173(6):779–791.
- Bijma, P., Muir, W. M., and Van Arendonk, J. A. (2007). Multilevel selection 1: Quantitative genetics of inheritance and response to selection. *Genetics*, 175(1):277–288.
- Birch, J. and Okasha, S. (2015). Kin selection and its critics. *BioScience*, 65(1):22–32.
- Blatrix, R. and Herbers, J. (2003). Coevolution between slave-making ants and their hosts: host specificity and geographical variation. *Mol. Ecol.*, 12(10):2809–2816.
- Blyth, C. R. (1972). On Simpson’s paradox and the sure-thing principle. *J. Am. Stat. Assoc.*, 67(338):364–366.
- Boomsma, J. J., Beekman, M., Cornwallis, C. K., Griffin, A. S., Holman, L., Hughes, W. O., Keller, L., Oldroyd, B. P., and Ratnieks, F. L. (2011). Only full-sibling families evolved eusociality. *Nature*, 471(7339):E4–E5.
- Boots, M., Best, A., Miller, M. R., and White, A. (2009). The role of ecological feedbacks in the evolution of host defence: what does theory tell us? *Phil. Trans. R. Soc. B*, 364(1513):27–36.
- Boots, M. and Bowers, R. G. (2004). The evolution of resistance through costly acquired immunity. *Proc. R. Soc. B*, 271(1540):715–723.
- Boots, M. and Haraguchi, Y. (1999). The evolution of costly resistance in host-parasite systems. *Am. Nat.*, 153(4):359–370.
- Bourke, A. F. (2011). The validity and value of inclusive fitness theory. *Proc. R. Soc. B*, 278(1723):3313–3320.
- Boyd, R. and Lorberbaum, J. P. (1987). No pure strategy is evolutionarily stable in the repeated Prisoner’s Dilemma game. *Nature*, 327:58–59.
- Brandt, M., Fischer-Blass, B., Heinze, J., and Foitzik, S. (2007). Population structure and the co-evolution between social parasites and their hosts. *Mol. Ecol.*, 16(10):2063–2078.
- Brandt, M. and Foitzik, S. (2004). Community context and specialization influence coevolution between a slavemaking ant and its hosts. *Ecology*, 85(11):2997–3009.

- Brandt, M., Foitzik, S., Fischer-Blass, B., and Heinze, J. (2005a). The coevolutionary dynamics of obligate ant social parasite systems – between prudence and antagonism. *Biol. Rev.*, 80(2):251–267.
- Brandt, M., Heinze, J., Schmitt, T., and Foitzik, S. (2005b). A chemical level in the coevolutionary arms race between an ant social parasite and its hosts. *J. Evol. Biol.*, 18(3):576–586.
- Bshary, R. and Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441(7096):975–978.
- Buschinger, A. (1986). Evolution of social parasitism in ants. *Trends Ecol. Evol.*, 1(6):155–160.
- Csilléry, K., Blum, M. G., Gaggiotti, O. E., and François, O. (2010). Approximate Bayesian computation (ABC) in practice. *Trends Ecol. Evol.*, 25(7):410–418.
- Csilléry, K., François, O., and Blum, M. G. (2012). abc: an R package for approximate Bayesian computation (ABC). *Methods Ecol. Evol.*, 3(3):475–479.
- Darwin, C. R. (1859). *The origin of species*. J. Murray, London.
- Davies, N., Bourke, A. F., and Brooke, M. d. L. (1989). Cuckoos and parasitic ants: interspecific brood parasitism as an evolutionary arms race. *Trends Ecol. Evol.*, 4(9):274–278.
- Davies, N. and Brooke, M. d. L. (1989). An experimental study of co-evolution between the cuckoo, *Cuculus canorus*, and its hosts. i. host egg discrimination. *J. Anim. Ecol.*, pages 207–224.
- Davies, N., Brooke, M. d. L., and Kacelnik, A. (1996). Recognition errors and probability of parasitism determine whether reed warblers should accept or reject mimetic cuckoo eggs. *Proc. R. Soc. B*, 263(1372):925–931.
- Dawkins, R. (1976). *The selfish gene*. Oxford University Press, Oxford.
- Dawkins, R. and Krebs, J. R. (1979). Arms races between and within species. *Proc. R. Soc. B*, 205(1161):489–511.
- Débarre, F., Hauert, C., and Doebeli, M. (2014). Social evolution in structured populations. *Nat. Commun.*, 5(3409):1–7.

BIBLIOGRAPHY

- Débarre, F., Lion, S., van Baalen, M., and Gandon, S. (2012). Evolution of host life-history traits in a spatially structured host-parasite system. *Am. Nat.*, 179(1):52–63.
- Diggle, S. P., Griffin, A. S., Campbell, G. S., and West, S. A. (2007). Cooperation and conflict in quorum-sensing bacterial populations. *Nature*, 450(7168):411–414.
- Dobrinevski, A. and Frey, E. (2012). Extinction in neutrally stable stochastic Lotka-Volterra models. *Phys. Rev. E*, 85(5):051903.
- Duncan, A. B., Fellous, S., and Kaltz, O. (2011). Reverse evolution: selection against costly resistance in disease-free microcosm populations of *Paramecium Caudatum*. *Evolution*, 65(12):3462–3474.
- Dunford, C. (1977). Kin selection for ground squirrel alarm calls. *Am. Nat.*, 111(980):782–785.
- Durrett, R. (1995). Ten lectures on particle systems. In Bernard, P., editor, *Lectures on Probability Theory*, volume 1608 of *Lecture Notes in Mathematics*, pages 97–201. Springer Berlin Heidelberg.
- Durrett, R. (2008). *Probability models for DNA sequence evolution*. Springer Science & Business Media, New York, 2nd edition.
- Durrett, R. and Levin, S. (1994). The importance of being discrete (and spatial). *Theor. Popul. Biol.*, 46(3):363–394.
- Emery, C. (1909). Über den Ursprung der dulotischen, parasitischen und myrmekophilen Ameisen. *Biol. Centralblt.*, 29:352–362.
- Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes: Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.
- Farrell, J. and Ware, R. (1989). Evolutionary stability in the repeated Prisoner’s Dilemma. *Theor. Popul. Biol.*, 36(2):161–166.
- Ferriere, R. and Michod, R. E. (2011). Inclusive fitness in evolution. *Nature*, 471(7339):E6–E8.
- Fischer, B. and Foitzik, S. (2004). Local co-adaptation leading to a geographical mosaic of coevolution in a social parasite system. *J. Evol. Biol.*, 17(5):1026–1034.

- Fischer-Blass, B., Heinze, J., and Foitzik, S. (2006). Microsatellite analysis reveals strong but differential impact of a social parasite on its two host species. *Mol. Ecol.*, 15(3):863–872.
- Fisher, R. A. (1958). *The genetical theory of natural selection*. Dover, London, 2nd edition.
- Fletcher, J. A. and Zwick, M. (2007). The evolution of altruism: Game theory in multilevel selection and inclusive fitness. *J. Theor. Biol.*, 245(1):26–36.
- Foitzik, S., Achenbach, A., and Brandt, M. (2009). Locally adapted social parasite affects density, social structure, and life history of its ant hosts. *Ecology*, 90(5):1195–1206.
- Foitzik, S., Backus, V. L., Trindl, A., and Herbers, J. M. (2004). Ecology of *Leptothorax* ants: impact of food, nest sites, and social parasites. *Behav. Ecol. Sociobiol.*, 55(5):484–493.
- Foitzik, S., DeHeer, C. J., Hunjan, D. N., and Herbers, J. M. (2001). Coevolution in host–parasite systems: behavioural strategies of slave–making ants and their hosts. *Proc. R. Soc. B*, 268(1472):1139–1146.
- Foitzik, S., Fischer, B., and Heinze, J. (2003). Arms races between social parasites and their hosts: geographic patterns of manipulation and resistance. *Behav. Ecol.*, 14(1):80–88.
- Foitzik, S. and Herbers, J. M. (2001a). Colony structure of a slavemaking ant. I. Intracolony relatedness, worker reproduction, and polydomy. *Evolution*, 55(2):307–315.
- Foitzik, S. and Herbers, J. M. (2001b). Colony structure of a slavemaking ant. II. Frequency of slave raids and impact on the host population. *Evolution*, 55(2):316–323.
- Frank, S. A. (2000). Specific and non-specific defense against parasitic attack. *J. Theor. Biol.*, 202(4):283–304.
- Fukuyo, M., Sasaki, A., and Kobayashi, I. (2012). Success of a suicidal defense strategy against infection in a structured habitat. *Sci. rep.*, 2(238):1–8.
- Gardner, A. (2008). The Price equation. *Current Biology*, 18(5):R198–R202.
- Gardner, A. (2015). The genetical theory of multilevel selection. *J. Evol. Biol.*, 28(2):305–319.
- Gardner, A. and West, S. A. (2010). Greenbeards. *Evolution*, 64(1):25–38.
- Gardner, A., West, S. A., and Wild, G. (2011). The genetical theory of kin selection. *J. Evol. Biol.*, 24(5):1020–1043.

BIBLIOGRAPHY

- Gillespie, J. (1974). Polymorphism in patchy environments. *Am. Nat.*, 108(960):145–151.
- Gladstone, D. E. (1981). Why there are no ant slave rebellions. *Am. Nat.*, 117(5):779–781.
- Goodnight, C. J. and Stevens, L. (1997). Experimental studies of group selection: what do they tell us about group selection in nature? *Am. Nat.*, 150(S1):s59–s79.
- Grafen, A. (1984). Natural selection, kin selection and group selection. In Krebs, J. R. and Davies, N. B., editors, *Behavioural ecology: an evolutionary approach*, pages 62–84. Blackwell Scientific Publications, Oxford.
- Grafen, A. (1985). A geometric view of relatedness. In Dawkins, R. and Ridley, M., editors, *Oxford surveys in evolutionary biology*, volume 2, pages 28–89. Oxford University Press, Oxford.
- Grafen, A. (2007). An inclusive fitness analysis of altruism on a cyclical network. *J. Evol. Biol.*, 20(6):2278–2283.
- Hamilton, W. D. (1963). The evolution of altruistic behavior. *Am. Nat.*, 97:354–356.
- Hamilton, W. D. (1964a). The genetical evolution of social behaviour. I. *J. Theor. Biol.*, 7(1):1–16.
- Hamilton, W. D. (1964b). The genetical evolution of social behaviour. II. *J. Theor. Biol.*, 7(1):17–52.
- Handel, A., Longini, I. M., and Antia, R. (2009). Antiviral resistance and the control of pandemic influenza: the roles of stochasticity, evolution and model details. *J. Theor. Biol.*, 256(1):117–125.
- Haraguchi, Y. and Sasaki, A. (2000). The evolution of parasite virulence and transmission rate in a spatially structured population. *J. Theor. Biol.*, 203(2):85–96.
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.
- Hare, J. F. and Alloway, T. M. (1987). Early learning and brood discrimination in leptothoracine ants (hymenoptera: Formicidae). *Anim. Behav.*, 35(6):1720–1724.
- Hauert, C. and Schuster, H. G. (1997). Effects of increasing the number of players and memory size in the iterated Prisoner’s Dilemma: a numerical approach. *Proc. R. Soc. B*, 264(1381):513–519.

- Heinze, J., Ortius, D., Kaib, M., and Hölldobler, B. (1994). Interspecific aggression in colonies of the slave-making ant *Harpagoxenus sublaevis*. *Behav. Ecol. Sociobiol.*, 35(2):75–83.
- Herbers, J. M. (1989). Community structure in north temperate ants: temporal and spatial variation. *Oecologia*, 81(2):201–211.
- Herbers, J. M. and Cunningham, M. (1983). Social organization in *Leptothorax longispinosus* Mayr. *Anim. Behav.*, 31(3):759–771.
- Herbers, J. M. and Foitzik, S. (2002). The ecology of slavemaking ants and their hosts in north temperate forests. *Ecology*, 83(1):148–163.
- Herbers, J. M. and Stuart, R. J. (1998). Patterns of reproduction in slave-making ants. *Proc. R. Soc. B*, 265(1399):875–887.
- Herre, E. A. and Wcislo, W. T. (2011). In defence of inclusive fitness theory. *Nature*, 471(7339):E8–E9.
- Hölldobler, B. and Wilson, E. O. (2009). *The superorganism: the beauty, elegance, and strangeness of insect societies*. WW Norton & Company, New York.
- Howard, R. W. and Blomquist, G. J. (1982). Chemical ecology and biochemistry of insect hydrocarbons. *Annu. Rev. Entomol.*, 27(1):149–172.
- Howard, R. W. and Blomquist, G. J. (2005). Ecological, behavioral, and biochemical aspects of insect hydrocarbons*. *Annu. Rev. Entomol.*, 50:371–393.
- Hutzenthaler, M. (2009). The virgin island model. *Electron. J. Probab.*, 14(39):1117–1161.
- Hutzenthaler, M. (2012). Interacting diffusions and trees of excursions: convergence and comparison. *Electron. J. Probab.*, 17(71):1–49.
- Hutzenthaler, M., Jordan, F., and Metzler, D. (2015). Altruistic defense traits in structured populations. arXiv:1505.02154v1.
- Hutzenthaler, M. and Wakolbinger, A. (2007). Ergodic behavior of locally regulated branching populations. *Ann. Appl. Probab.*, 17(2):474–501.
- Johnson, C. A. and Herbers, J. M. (2006). Impact of parasite sympatry on the geographic mosaic of coevolution. *Ecology*, 87(2):382–394.

BIBLIOGRAPHY

- Karlin, S. and Taylor, H. M. (1981). *A second course in stochastic processes*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York.
- Klenke, A. (2008). *Probability theory*. Universitext. Springer-Verlag London Ltd., London. A comprehensive course, Translated from the 2006 German original.
- Klenke, A. and Mytnik, L. (2012). Infinite rate mutually catalytic branching in infinitely many colonies: construction, characterization and convergence. *Probab. Theory Related Fields*, 154(3-4):533–584.
- Konrad, M., Pamminger, T., and Foitzik, S. (2012). Two pathways ensuring social harmony. *Naturwissenschaften*, 99(8):627–636.
- Kurtz, T. G. (1992). Averaging for martingale problems and stochastic approximation. In *Applied stochastic analysis (New Brunswick, NJ, 1991)*, volume 177 of *Lecture Notes in Control and Inform. Sci.*, pages 186–209. Springer, Berlin.
- Labrie, S. J., Samson, J. E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nature Rev. Microbiol.*, 8(5):317–327.
- Lahav, S., Soroker, V., Hefetz, A., and Vander Meer, R. K. (1999). Direct behavioral evidence for hydrocarbons as ant recognition discriminators. *Naturwissenschaften*, 86(5):246–249.
- Lande, R. (1976). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.*, 26(3):221–235.
- Langmore, N. E., Hunt, S., and Kilner, R. M. (2003). Escalation of a coevolutionary arms race through host rejection of brood parasitic young. *Nature*, 422(6928):157–160.
- Lehmann, L., Keller, L., West, S., and Roze, D. (2007). Group selection and kin selection: two concepts but one process. *Proc. Natl. Acad. Sci. USA*, 104(16):6736–6739.
- Lehmann, L. and Rousset, F. (2010). How life history and demography promote or inhibit the evolution of helping behaviours. *Phil. Trans. R. Soc. B*, 365(1553):2599–2617.
- Lenormand, T., Roze, D., and Rousset, F. (2009). Stochasticity in evolution. *Trends Ecol. Evol.*, 24(3):157–165.
- Levene, H. (1953). Genetic equilibrium when more than one ecological niche is available. *Am. Nat.*, pages 331–333.

- Lieberman, E., Hauert, C., and Nowak, M. A. (2005). Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316.
- Lion, S., Jansen, V. A., and Day, T. (2011). Evolution in structured populations: beyond the kin versus group debate. *Trends Ecol. Evol.*, 26(4):193–201.
- Lotka, A. J. (1920). Undamped oscillations derived from the law of mass action. *J. Am. Chem. Soc.*, 42(8):1595–1599.
- Marshall, J. A. (2011). Group selection and kin selection: formally equivalent approaches. *Trends Ecol. Evol.*, 26(7):325–332.
- Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, 201:1145–1147.
- McGregor, P. K., editor (2005). *Animal communication networks*. Cambridge University Press, Cambridge, UK.
- Metzler, D., Jordan, F., Pamminger, T., and Foitzik, S. (2015). The influence of space and time on the evolution of altruistic defense: the case of ant slave rebellion. Unpublished manuscript.
- Milinski, M. (1987). Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, 325(6103):433–435.
- Miller, M. R., White, A., and Boots, M. (2007). Host life span and the evolution of resistance characteristics. *Evolution*, 61(1):2–14.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563.
- Nowak, M. A. and May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829.
- Nowak, M. A., Tarnita, C. E., and Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466(7310):1057–1062.
- Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505.
- Okasha, S. (2006). *Evolution and the levels of selection*. Oxford University Press, Oxford.
- Okasha, S. (2010). Altruism researchers must cooperate. *Nature*, 467(7316):653–655.

BIBLIOGRAPHY

- Øksendal, B. (1998). *Stochastic differential equations – An introduction with applications*. Universitext. Springer-Verlag, Berlin, fifth edition.
- Orlove, M. and Wood, C. L. (1978). Coefficients of relationship and coefficients of relatedness in kin selection: a covariance form for the RHO formula. *J. Theor. Biol.*, 73(4):679–686.
- Ortega, C. P. (1998). *Cowbirds and other brood parasites*. University of Arizona Press, Tucson.
- Pamminger, T., Foitzik, S., Metzler, D., and Pennings, P. S. (2014). Oh sister, where art thou? Spatial population structure and the evolution of an altruistic defence trait. *J. Evol. Biol.*, 27(11):2443–2456.
- Pamminger, T., Leingärtner, A., Achenbach, A., Kleeberg, I., Pennings, P. S., and Foitzik, S. (2013). Geographic distribution of the anti-parasite trait “slave rebellion”. *Evol. Ecol.*, 27(1):39–49.
- Pamminger, T., Modlmeier, A. P., Suetter, S., Pennings, P. S., and Foitzik, S. (2012). Raiders from the sky: slavemaker founding queens select for aggressive host colonies. *Biol. Lett.*, 8(5):748–750.
- Pamminger, T., Scharf, I., Pennings, P. S., and Foitzik, S. (2011). Increased host aggression as an induced defense against slave-making ants. *Behav. Ecol.*, 22(2):255–260.
- Parvinen, K. (2010). Adaptive dynamics of cooperation may prevent the coexistence of defectors and cooperators and even cause extinction. *Proc. R. Soc. B*, 277:2493–2501.
- Parvinen, K., Dieckmann, U., Gyllenberg, M., and Metz, J. A. (2003). Evolution of dispersal in metapopulations with local density dependence and demographic stochasticity. *J. Evol. Biol.*, 16(1):143–153.
- Pennings, P. S., Achenbach, A., and Foitzik, S. (2011). Similar evolutionary potentials in an obligate ant parasite and its two host species. *J. Evol. Biol.*, 24(4):871–886.
- Pitman, J. and Yor, M. (1982). A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete*, 59(4):425–457.
- Platt, T. G. and Bever, J. D. (2009). Kin competition and the evolution of cooperation. *Trends Ecol. Evol.*, 24(7):370–377.

- Pohl, S. and Foitzik, S. (2011). Slave-making ants prefer larger, better defended host colonies. *Anim. Behav.*, 81(1):61–68.
- Price, G. R. (1970). Selection and covariance. *Nature*, 227:520–21.
- Queller, D. and Goodnight, K. (1989). Estimation of genetic relatedness using allozyme data. *Evolution*, 43:258–275.
- Queller, D. C. (1992). Quantitative genetics, inclusive fitness, and group selection. *American Naturalist*, 139(3):540–558.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rauch, E. M., Sayama, H., and Bar-Yam, Y. (2002). Relationship between measures of fitness and time scale in evolution. *Phys. Rev. Lett.*, 88(22):228101.
- Rauch, E. M., Sayama, H., and Bar-Yam, Y. (2003). Dynamics and genealogy of strains in spatially extended host–pathogen models. *J. Theor. Biol.*, 221(4):655–664.
- Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge. Itô calculus, Reprint of the second (1994) edition.
- Rousset, F. and Lion, S. (2011). Much ado about nothing: Nowak et al.’s charge against inclusive fitness theory. *J. Evol. Biol.*, 24(6):1386–1392.
- Scharf, I., Bauer, S., Fischer-Blass, B., and Foitzik, S. (2011). Impact of a social parasite on ant host populations depends on host species, habitat and year. *Biol. J. Linnean Soc.*, 103(3):559–570.
- Sherman, P. W. (1977). Nepotism and the evolution of alarm calls. *Science*, 197(4310):1246–1253.
- Shorter, J. and Rueppell, O. (2012). A review on self-destructive defense behaviors in social insects. *Insectes Soc.*, 59(1):1–10.
- Sibly, R. M. and Curnow, R. (2011). Selfishness and altruism can coexist when help is subject to diminishing returns. *Heredity*, 107(2):167–173.
- Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *J. R. Stat. Soc. B*, 13(2):238–241.

BIBLIOGRAPHY

- Slatkin, M. and Wade, M. (1978). Group selection on a quantitative character. *Proc. Nat. Acad. Sci. USA*, 75(7):3531–3534.
- Sober, E. and Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Harvard University Press, Cambridge, MA.
- Strassmann, J. E., Page Jr, R. E., Robinson, G. E., and Seeley, T. D. (2011). Kin selection and eusociality. *Nature*, 471(7339):E5–E6.
- Stuart, R. J. and Alloway, T. M. (1985). Behavioural evolution and domestic degeneration in obligatory slave-making ants (Hymenoptera: Formicidae: Leptothoracini). *Anim. Behav.*, 33(4):1080–1088.
- Taylor, P. (1992). Altruism in viscous populations—an inclusive fitness model. *Evol. Ecol.*, 6(4):352–356.
- Taylor, P. D., Day, T., and Wild, G. (2007). Evolution of cooperation in a finite homogeneous graph. *Nature*, 447(7143):469–472.
- Taylor, P. D. and Frank, S. A. (1996). How to make a kin selection model. *J. Theor. Biol.*, 180(1):27–37.
- Thompson, J. N. (1999). Specific hypotheses on the geographic mosaic of coevolution. *Am. Nat.*, 153(S5):S1–S14.
- Traulsen, A. (2010). Mathematics of kin- and group-selection: formally equivalent? *Evolution*, 64(2):316–323.
- Traulsen, A. and Nowak, M. A. (2006). Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. USA*, 103(29):10952–10955.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.*, 46(1):35–57.
- Turner, P. E. and Chao, L. (1999). Prisoner’s dilemma in an RNA virus. *Nature*, 398(6726):441–443.
- Uyenoyama, M. K. (1979). Evolution of altruism under group selection in large and small populations in fluctuating environments. *Theor. Popul. Biol.*, 15(1):58–85.
- Van Baalen, M. and Rand, D. A. (1998). The unit of selection in viscous populations and the evolution of altruism. *J. Theor. Biol.*, 193(4):631–648.

- Van Valen, L. (1971). Group selection and the evolution of dispersal. *Evolution*, 25(4):591–598.
- Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1:1–30.
- Van Valen, L. (1974). Molecular evolution as predicted by natural selection. *J. Mol. Evol.*, 3(2):89–101.
- Van Veelen, M. (2009). Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. *J. Theor. Biol.*, 259(3):589–600.
- Van Veelen, M., García, J., Sabelis, M. W., and Egas, M. (2012). Group selection and inclusive fitness are not equivalent; the Price equation vs. models and statistics. *J. Theor. Biol.*, 299(0):64–80.
- Van Veelen, M., Luo, S., and Simon, B. (2014). A simple model of group selection that cannot be analyzed with inclusive fitness. *J. Theor. Biol.*, 360(0):279–289.
- Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*, 118:558–560.
- Wade, M. J. (1976). Group selections among laboratory populations of tribolium. *Proc. Natl. Acad. Sci. USA*, 73(12):4604–4607.
- Wade, M. J. (1978). A critical review of the models of group selection. *Q. Rev. Biol.*, 53(2):101–114.
- Wade, M. J. (1980). Kin selection: its components. *Science*, 210(4470):665–667.
- Wade, M. J. (1982). Evolution of interference competition by individual, family, and group selection. *Proc. Natl. Acad. Sci. USA*, 79(11):3575–3578.
- Ward, P. S. (2014). The phylogeny and evolution of ants. *Annu. Rev. Ecol. Evol. Syst.*, 45:23–43.
- Wesson, L. G. (1939). Contributions to the natural history of *Harpagoxenus americanus* Emery (Hymenoptera: Formicidae). *Trans. Am. Entomol. Soc.*, 65(2):97–122.
- West, S. A., Griffin, A. S., and Gardner, A. (2007a). Evolutionary explanations for cooperation. *Curr. Biol.*, 17(16):R661–R672.
- West, S. A., Griffin, A. S., and Gardner, A. (2007b). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.*, 20(2):415–432.

BIBLIOGRAPHY

- Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, 308(5955):181–184.
- Wilkinson, G. S. (1988). Reciprocal altruism in bats and other mammals. *Ethol. Sociobiol.*, 9(2):85–100.
- Wilkinson, G. S. (1990). Food sharing in vampire bats. *Sci. Am.*, 262(2):76–82.
- Williams, P., Winzer, K., Chan, W. C., and Camara, M. (2007). Look who’s talking: communication and quorum sensing in the bacterial world. *Phil. Trans. R. Soc. B*, 362(1483):1119–1134.
- Wilson, D. S. (1975). A theory of group selection. *Proc. Natl. Acad. Sci. USA*, 72(1):143–146.
- Wilson, D. S. (1979). Structured demes and trait-group variation. *Am. Nat.*, 113(4):606–610.
- Wilson, D. S., Pollock, G., and Dugatkin, L. A. (1992). Can altruism evolve in purely viscous populations? *Evol. Ecol.*, 6(4):331–341.
- Wirsching, G. J. (2006). *Gewöhnliche Differentialgleichungen: eine Einführung mit Beispielen, Aufgaben und Musterlösungen*. Teubner, Wiesbaden.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2):97–159.
- Wright, S. (1945). Tempo and mode in evolution: a critical review. *Ecology*, 26(4):415–419.
- Wright, S. (1948). Genetics of populations. *Encyclopedia Britannica*, 10:111–112.
- Yamada, T. and Watanabe, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167.
- Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, 2(2):121–134.

Acknowledgments

First of all, I would like to thank Dirk Metzler and Martin Hutzenthaler. Thank you both for your patience and for being so generous with your help and advice throughout the years. I am extremely grateful that you trusted me with this project and always made the time to provide all the support and feedback that I could have wished for.

Next, I would like to thank Niels Dingemanse, Christian Leibold, and Stephan Hutter for joining my thesis advisory committee and taking the time to meet with me and giving valuable and much appreciated feedback.

Furthermore, I am thankful to the wonderful group of Evolutionary Biology. I would especially like to thank my colleagues Paul Staab, Dimitrios Lyras, and Meike Wittmann for many helpful discussions.

I am very grateful for the funding from Deutsche Forschungsgemeinschaft in the Priority Programme *Probabilistic Structures in Evolution* (SPP 1590).